A SELECT CATALOG
of
LANGUAGE UNIVERSALS

EUROLINGUA
P.O. Box 101
Bloomington, Indiana 47402-0101
1988

A SELECT CATALOG OF LANGUAGE UNIVERSALS

Bibliotheca Nostratica

Edited

by

GYULA DÉCSY

Volume 8

GYULA DÉCSY

A SELECT CATALOG OF LANGUAGE UNIVERSALS

EUROLINGUA

Bloomington, Indiana

1988

GYULA DÉCSY

A SELECT CATALOG OF LANGUAGE UNIVERSALS

EUROLINGUA

Bloomington, Indiana

1987

ISSN 0342-4871 ISBN 931922-29-1

All Rights Reserved

No part of this book may be reproduced in any form by print (including xerox, photoprint, microfilm) or by any other means (including audio) without prior written permission of the author.

© Gyula Décsy 1988, Bloomington, Indiana

CONTENTS

INTRODUCTORY REMARKS 9

1. PRESOUNDSEQUENTIAL LANGUAGE UNIVERSALS 2

Preliminaries: Prehuman/Subhuman 21 – Priority of Physical Contact 24 – Non-Tactile-Sensical Contacts 25 – E(lec)tric and Anorganic: Gravity 25 – Chemistry: Taste/Smell 26 – Audio/Hearing 27 – Vision/Sight 27 – Signal Noise Production 29 – Gestures 28 – Pulmonary Vocalization 31 – Concept and Concept Pressure 32 – Concept and Sign Types 33 – Sign System Types 35 – Brain as Speech Organ 35 – Bit Storage in the Human Brain 38 – Language Acquisition and Language Attrition 40 – Recapitulationism 41.

2. SOUNDSEQUENTIAL LANGUAGE UNIVERSALS 43

Generic Tenets 43 – Duality of Patterning is Possible only in Soundsequence 44 – Identical Deep Structure in the Languages of the World 46 – Evolutionary/Hierarchic Buildup of Grammar 45 – Three Basic Types of Linguistic Change 47

2.1 PHONETICS/PHONOLOGY 48

Sounds/Phonemes 48-49 – Vowels 52 – Vowel systems 53 – Long Vowels 55 – Diphthongs 56 – Consonants 57 – The Consonant /h/ 57 – Glottal Stop 58 – Stops 58 – Sonority Sets 59 – Fricatives 59 – Affricates 60 – Liquids (Laterals and Vibrants) 61 – Retroflexes 61 – Implosives 61 – Consonant Accessorials 62 – Long Consonants 62 – Nasals 63 – Increasing Phoneme Numbers 64 – Sound Icocinity/Toconism 65 – Evolution and Markedness 67 – Frequency of Occurence 69 – Homophony/Opacity/Transparence 70 – Syllables and Liaison 70 – Pause as a Prosodeme 74 – Register 74 - Versification 74 – Mora 73 – Syllable Tones 73 – Sound Change 76 Alternations 79 – Vowel Harmony 80 – Synharmony 81 – Aesthetic Value of Sounds and Soundsequences 81.

2.2 MORPHOLOGY 82

Parts of the Speech/Word Classes 84 – Word 84 – Gender; Cass 85 – Grammemes/Affixes/Formants 86 – Number 87-88 – Personal Pronouns/Category of Person 89 – Derivation/Word Production 90 – Diminutives and Their Derivational Priority 91 – Paradigms/Inflectional Sets 92 – Cases/Declension 92 – Verbalnouns 94 – Possessive Inflection 95 – Verb/Conjugation 95 – Mood 96 – Secondary Moods 96-97 – Aktionsart 97 – Continuous Forms 97 – Experience Form/Reference 97 – Synonyms 101 – Homonyms and Phoricity; Situative Reference/Context 101 – Tense and Time 98 – Word Length 100 – Honorifics/Deferentials 102.

2.3 SYNTAX......103
Predicate and Subject 105 – Object 105 – Object 106 – Adverbials 106 –Word Order 107 – Sentence Tone: Interrogatives 109 – Compound/Multiple Sentences and Conjunctions 111 – Affirmation/ Negation 110 – Ergativity 112.

2.4 SEMANTICS......113
Paleo-Sememes 113 – Metaphor; Idioms 114 – Species 116 – Inexactness/ Prevarication 119 – Emotional/Concrete/Abstract 119 – Some Archaic Etymons 120 – habeo 121.

2.5 VOCABULARY/LEXEMICS......122
Color Terms 123 – Lexemic Transparence/Lexemic Opacity 124 – Negation 124-125 – Synaesthesia 125 – Smell and Taste125 – Word Retention Stability 126 – Borrowings: Words and Calques 127.

2.6. OLINGUISTICS......128
Geolinguistics 128 – Standard Language/Acrolect 129 – Dialect 131 – Macro-Sociolinguistics 131 – Script and its Decoding 130 – Language Comparison 133 – Linguistic Futurology 135.

REFERENCES AND ABBREVIATIONS......137

Diagrams:
Vocal Cube (Vokalklotz)......53
Sonority Set......59

INTRODUCTORY REMARKS

The study of Language Universals on a large scale was initiated by Joseph Greenberg and his associates in the 60s, i.e. hardly more than a quarter century ago. Their organizational efforts resulted in the Stanford Universal Project on Language Universals, which began its activities in October 1967 and brought them to an end in August 1976 (S 1.VI). The ten-year project was one of the largest institutionalized enterprises of linguistic research in the U.S.A. in recent decades; its staff included thirty-two part-time or full-time linguists. The final results of the Stanford research were published in a four-volume work under the title Universals of Human Language (edited by J. H. Greenberg; Associate Editors were Ch. A. Ferguson and E. A. Moravcsik. See in the list of References and Abbreviations under S below). As a visiting scholar, Hansjakob Seiler (University of Köln) collaborated with the program at Stanford; he then set up his own research program, known as the Cologne Project on Language Universals (Kölner Universalienprojekt), in Germany. The most important findings of the Cologne research team were disseminated by several voluminous publications (see G, HS, K, and SB in the References and Abbreviations below). The members of the Stanford Project were mostly Americans, while the Cologne team consisted mainly of young German scholars. Some researchers from Eastern Europe (Gamkrelidze, L. Dezső) cooperated with the Stanford and Cologne programs; however, a somewhat more systematic research on language universals in the eastern countries was carried out in the 60s and 70s in Russia. The Russian linguists-universalists (stimulated mainly by Greenberg's publications) were fascinated first of all by the Stanford Project; they published some chapters on universals in books dealing chiefly with "structural typology" (a favorite topic of Russian linguists – see Uspenskij 1965, Nauka 1974, Khrakovskij 1977, Sharadzenidze 1982). At a time when linguists in America were attracted to abstract theoretical problems, the Stanford and Cologne projects exhibited an impressive and solid data-orientedness in their methods.

Ten years have passed since the conclusion both of the Stanford and the Cologne language universal projects. Since then, the strongly theory-oriented books of the prolific English-American scholar Comrie (1981 and BCD 1984; see in References and Abbreviations below),an impressive symposium at the University of Wisconsin at Milwaukee (coordinated by E. Moravcsik – see UAJb 58.152-153, 1986), and some planned publications of M. Ruhlen (as Volume 3 of his Guide to the World's Languages) were major remarkable events in research on language universals. However, there are no large institutionalized research projects on the topic in progress any more. This suggests that, according to our leading linguists, the problems of language universals have already been resolved or the research on the topic is without prospect (see G. Diffloth's skeptical views: UAJb 58.153). The present publication (in its entirety the result of individual research), on the contrary, intends to demonstrate that there is still room for new approaches and ideas for research in the field. It also calls attention to methods and principles which have remained (certainly inadvertently) outside the interest of previous institutionalized research.

There is no acceptable definition of the language universal. A self-contained view would be that a language universal (LU) is a linguistic feature common to all present-day and extinct natural languages (without chronological or geographical restriction) everywhere in the world. This panchronic-global tenet is kept, apparently, by the suggestion from Generative Grammar that there is one innate language structure in the world which has existed in the mind of all human beings from the beginning of times; it exists there at present and will exist there unchanged to the end of times. The Stanford project never shared this somewhat dogmatic perception of language universals and concentrated (at least at its conception) on the collection of a large number of empirical universals by means of impressive attempts of well-trained individuals at cross-linguistic (global) language comparisons with regard to certain linguistic characteristics. Seiler's Cologne program group consciously liberalized the concept of the language universal (HS 11-25, especially 20); their representatives saw the language universals as hardly more than accurate linguistic statements concerning invariants with presumably global validity made on a certain (mostly limited) number of languages. The Cologne team concept eased the conceptual restrictions; however, their outcomes exhibit less coherence than the more broadly based and fact-oriented results of Stanford. In the final publication of their findings, both the Stanford and the Cologne universalists used then (following trends in linguistic studies of the 60s and 70s) all chances for theory building which were offered by their material collections. Unfortunately, the preponderance of theory often obscured a considerable portion of the facts which were collected (or, at least, distracted attention from them). The present publication, besides its committed data-orientedness, tries to offer fresh illustrative materials drawn mainly from areas which were not always readily accessible (or attractive enough) to earlier researchers on language universals (Uralic, Slavic, and Siberian languages: Language Origins Research, Protolinguistics, Pre-Protolinguistics).

The definitions of language universals have revealed a certain uncertainity from the beginning. The term was introduced, as is well known, by Aginsky & Aginsky in 1948 (Word 4.168-172). The further history of the concept was impressively outlined by Ch. Ferguson (S 1.9-32). There is no place in this book for exact quotations of the almost one hundred excellent attempts at definitions (many are to be found in Greenberg s, Ferguson s, and Hockett's contributions in UL and S). However, I find important to print here the following small sampling of current terms (most of which are self-explanatory) used mainly by earlier researchers:

absolute universal (S 13)
absolute implicational universal (Comrie 1981.203)
accidental universals (Hockett UL 6, J. Moravcsik S 1.18)
almost absolute implicational universal (Comrie 1981.203)
analytical universals (J. Moravcsik S 1.18)
definitional universals (J. Moravcsik 1967, S 18)
detail universals (e.g. Greenberg's 54 generalizations in S 3.253-289)
essential universals (Hockett UL 6)
global universals (in this book)
implicational universals (Jakobson in 1958 in Oslo, S. 1.13)
intragenetic diachronical universal (Greenberg in S 3.274)
language origins research (LOR) universals (in this book)
linguistic universals (Katz/Postal 1964, a preferred TG/GG term, S 1.9)
near universals (S 1.13)
nondefinitional universals (J. Moravcsik 1967, S. 18)
non-human language universals (in this book)
ontogenetic language universal
pre-human language universals (in this book)
putative universals (Hockett UL 1, S 1.21)
quasi universals (Hoenigswald UL 49)
phylogenetic language universals
self-evident universals
statistical universals (J. Moravcsik 1967, S. 18)
strong universals (S 1.18)
subhuman language universals
synthetic universals (Moravcsik S 1.18)
tentative language universals (E. Moravcsik S 3.297)
trivial universals (J. Moravcsik 1967, S 18)
universals of human language (title of S)
weak universals (S 1.18)

complex invariant connection (Seiler HS 20)
cross-linguistic features
cross-linguistic generalizations (Greenberg S1.33)
crosslinguistic (no hyphen !) consistencies (UAJb 58.153)
general tendencies (Downing S 4.381)
implicational tendencies (Downing S 4.381)
invariants (Seiler HS 20)
omnipresent/panchronic (St. Ullmann UL 220)

Not included in this list are the so-called grammar-oriented classificatory terms (phonetic/phonological universal, morphological universal, grammatical universal, syntactic universal, etc.).

"Language universal" is not a very good term. It has been borrowed from the universal (Latin universale/universalia) concept of philosophy where it is connected with the name of A. M. S. Boethius (ca. A.D. 480-524), G. de Champeaux (d. 1211), and Thomas Aquinas (1225-1274). Language Universal (LU) implies that a given linguistic feature is of absolute validity not only on our Earth but also with regard to possible communication systems (languages) of intelligent beings on extraterrestrial-celestial bodies. If we want to remain in a realistic terrestrial/tellurian frame we should speak of "globals" ("globale/globalia", eventually "globemes"). If there are really intelligent beings on other celestial bodies, their anatomy (and, their entire body construction) may be completely different from that of the human beings on our Earth; if they communicate (which they certainly do), their signal systems may utilize elements which are unimaginably different from those of earthly citizens. It is, therefore, pretentious to extend earthbound language universals to the unknown huge parts of the universe. However, the term "language universal" is so commonly accepted that it is impossible to replace it with another, more suitable, name.

No doubt that there are such absolute language universals somewhere in the "deep structure" of the linguistic consciousness of mankind. However, they often escape emprical observation. I quote the elegant but rather skeptical formulation of St. Ullmann:

> It can never be proven conclusively that a phenomenon is "omnipresent" (panchronic), i.e. that "they exist in every language at any stage in its development" (Ullmann UL 220).

Nevertheless, if a phenomenon is really a language universal, then it is present in every *single* language. Yet we are unable to identify such absolute universals by the study of only one language. We have to check all languages of the world with regard to the presence of a certain feature (prospective universal) if we want to identify it even in just one single language. Language universal research is, therefore, the most comprehensive form of comparative linguistics. Needless to say, it is impossible to go through the grammar and vocabulary of three thousand or even twelve thousand "languages" of past and present (with an estimated 1.5 million individual data in every one of them) and to prove in this way that a certain linguistic phenomenon really existed or exists in all of them. Therefore, Ullmann's warning is fully justified. Nonetheless, we should try to find the cross-linguistic absolute universals. But, certainly, every linguist engaged in this type of research will deviate and compromise after a certain number of fruitless attempts: he will instead seek fast results by the convenient methods of "deduction" or generalization based on observations made on a small number of languages. He will then state that if there is a phenomenon which exists (provenly) in 100 languages, then it must (should, ought to or does) exist also in the remaining 2,700 or 11,900 languages (according to which extreme number the linguist accepts as the maximum for really existing/existed languages of the World). Or he begins to set up self-evident implicational universals which are probably not very insightful – statements like: every language has vowels; every language has at least one nasal consonant; if a language has a ts, it also has a t and s. Or he begins to labor on theoretical construction in the hope that in the end the theory is going to solve those problems of language universals which can not be resolved through empirical investigations. This last course has been followed by a large number of meritorious linguists doing work on language universals during the 60s and 70s.

Hockett stated that "if some feature is indeed universal, then it is taxonomically irrelevant" (LU 4; taxonomy means for Hockett approximately the same as typology). This statement suggests that it is not meaningful to handle universals and typology together. It is the view of this author that typology is one of the most confusing concepts of modern linguistics (he would prefer taxonomy or "language characteristic studies" – GLC 116-117). As is well known, Russian linguists consistently connect language universal research with various versions of typological studies (their principle is followed by B. Comrie in America). The preferred typology of the Russians is structural typology, a somewhat remodeled research technique of the Czech linguist Vl. Skalička. Indeed, both language universal research and typologies deal with characteristic features (properties) of languages (the comparative-historical method also does this). However, the object and the classification as well as the purpose of these various research techniques are completely different. The universalists deal with all languages of the world (as a part of a kind of global linguistics); the structural typologists study only smaller language groups which represent a certain grammatical type (agglutinative, synthetic/inflective, analytic, incorporating, stem isolating; see Dezső's excellent summary in HS 150-151) without any areal or other delineation. It is remarkable that the Russian universalists and typologists confuse "structural typology" with those branches of linguistic studies which are known as "areal typology" or "language alliance" (Sprachbund): they study mainly the area of the indigenous languages of East Eurasia, emphasizing the common features in their development (mainly the integrating influence of the Great Russian language in them) during the last several centuries. Even if it is certainly not easy to separate the areal features from the universals, language universal research should try to do it. It is methodologically wrong to handle and classify universals together with structural and areal features. They should be kept apart both in the research method and in the systematization of results.

The Stanford project's final findings were published under the title "Universals of *Human* Language" (italics mine, Gy. D.). The term *human* in it implies that there are *nonhuman* languages too, and that these nonhuman languages also have universals. Further, it implies that there are "prehuman" (or "subhuman") languages and prehuman language universals. All this makes a consistent separation of language and speech (parole) from each other a methodological necessity, emphasizing that language is something more comprehensive (universal) than speech. Language is not exclusively human, but speech (in its soundsequential realization) is human-species-specific (only-human). This is, of course, not a new idea, as it has already been an uncontroversial methodical principle since de Saussure's langue/parole (retermed by the TG as competence/performance) in linguistic studies. What is new here, however, is that the concept of language now comprises the communication systems of all (subhuman or prehuman) species using sign systems. The primates, the birds, the frogs all have a certain "language" according to this new interpretation (especially common in modern semiotic studies and in ethology). What the animals do not have, however, is "speech" ("parole" according to de Saussure's term). We can not emphasize often enough that not language, but speech is "only-human". Actually, we should speak about "speech universals" when we have "human universals" in mind. All this strengthens the position of "speech" in linguistic studies, which became somewhat peripheral during the last decade or so. At the present state of terminological convention it is, however, almost impossible to replace the term "(human) language universals" with "speech universals". The best solution is, perhaps, the one we find in the title of the Stanford publication: "language universal" should be used with the attribute "human", or, when without it, it should mean simply "human language universal". The enormous superiority of the human language (compared with the prehuman "languages") lies in the timbric sound, in the soundsequence and the extremely well organized syntax. These then are the areas which are the traditional subject of investigations on (human) language universals. However, there are a large number of prehuman and subhuman language universals, and they are all contained in some form in the human language universals. Prehuman/subhuman means, therefore (in this publication), first of all an evolutionary stage and not a feature type limited exclusively to the different forms of animal communication ("language of animals"). In other words, a prehuman or subhuman language universal is a human language universal which existed already in the prehuman and subhuman evolutionary/developmental stages.

A specific purpose of this publication is also to apply the principles of evolutionary theory to the research on language universals. Diachrony, and especially language prehistory (Language Origins Research), did not belong to the favorite research areas of either the Stanford or the Cologne projects. Remarkably, the (at least in principle or in their ideology) strongly Darwinistic Russian universalists (and linguists generally) disregard the results of modern evolution theory and diachrony too. A biased "antihistorism" has been a fashionable trend in linguistics since the takeover of structuralism in Europe and America in the 40s and 50s. To deal with the time aspect of language was regarded here and there as a kind of reactionism signaling inability to be "modern" and progressive in method (or even in political thinking). This attitude had changed somewhat by the end of the 70s. It is not difficult to predict that linguists of the 80s and 90s will have a lot to do in order to make good what the formalistic oriented generations of the last three decades omitted in their investigations with regard to the evolution and prehistory of language. A global phenomenon like language universal necessarily has a deeper time aspect than does a partial (genetic, areal, structural, cultural) segment of the world language scene. Anticipating this trends, in this book a large number of universals with "diachronic setting" were added, utilizing the research results of Long Range Language Comparison (LRLC), Language Origins Research (LOR) and, of course, also Protolinguistics (especially Indo-European and Uralic) as well as Pre-Protolinguistics (for the concepts see GLC). Evolutionary principles were claimed by Helmut Gipper from the Cologne language universal team, who stated: [It is] "a correct guiding-principle that, as in the biological realm in general, linguistic development leads from the simple to the complex" (K 545). We should not forget that "the arsenal of genetics is enlarged primarily by means of qualitative, not quantitative, changes in the program. Evolution therefore means primarily an increase in complexity" (Gipper K 542). However, it is often difficult to determine what is simple and what is complex. The analytic language system, for example, is simple (according to *opinio communis*), but its operational rules are more complex than those of the "complicated" synthetic language system. Generally, we suppose that the oldest language structures were analytic (with simple rules), and that they then became synthetic (more complex in form but still simple in operational rules). Then, in the next evolutionary stage, the complex synthetic language type with complex operational rules followed. A new analytic language type developed from a synthetic one (like Chinese or English) is then complicated both in structure and in operational rules. In any case, these modern analytic languages never return to the supposed simplicity of the prehistoric (amorphic/rulefree) language types.

I estimate that about 80 percent of the universals offered in this book (out of approximately 600) are new or basically reformulated. It is, as the title explicitly stated, only a select catalog. The selection is arbitrary and reflects some preferences of the author (especially for diachrony). From previous sorces, only such universals were included which appeared to be near-absolute, important and (in view of this author) ponderous for further research. These previous sources are, first of all, the Stanford and the Cologne projects' publications. I tried to proceed in the arrangement of the material from the simple to the complex, from the prehuman simple (natural) to the highly complicated phonemic, morphological, syntactic, vocabulary and "olinguistic" universals.

I mention separately that I also included a large number of statistical universals, mainly on the basis of Ruhlen's excellent publication in the Greenberg-Festschrift (see Ruhlen 1977 in the List of References and Abbreviations below). To state that 48 percent of the languages of the World have long vowels (and 52 percent do not have them – see Ruhlen 1977.150) is in my opinion a very important universal which is cross-linguistically of high significance, affecting practically (against J. Moravcsik as quoted in S 1.18) every single language of the world (either because it possesses or because it lacks the feature). Such statistical information (which was found and formulated with hard linguistic work) belong under any circumstances into the catalog of language universals.

Some commonly known and often quoted universals were, on the other hand, left out. Such are, among others, the popular and widely used "sentence word order" universals of Greenberg (SOV, SVO, VSO; see UL 110-111, numbers 1-25). I must confess that I have strong reservations with regard to this type of "universals". First of all, the V should be replaced in the symbols by P (in the sentence, we would have Predicate instead of "Verb"; "Verb" is an exclusive morphological concept). Second, there is no language in which we would not find counterexamples ("Gegenbeispiele") against Greenberg's rules. To state that the symbols SOV, SVO, VSO represent only general trends or neutral (basic, high frequency, enunciative, etc.) word order types means to circumvent the problem. There are, of course, clear cases of word order regularities, but only within the smaller units of the sentence (like attributive or prepositional constructions, postpositions, appositions; place of the predicate if fixed, e.g., at the end of the sentence); but there is no consistent word order with regard to the sequence of the parts of the sentence Subject/Predicate/Object in the languages of the world. In inflective (synthetic) languages, e.g., the subject is often "omitted" or included in the verbal ending (Latin video patrem venire; no ego is necessary). In many languages there is no genitive, and therefore, no genitive attribute or genitive constructions exist in them. In incorporating language types, the object may be expressed also by the verbal ending of the predicate (Hungarian látlak 'I see you'). How can we assign a fixed place in the SOV/SVO/VSO symbols for those parts of the sentence which do not exist as separate words ? See also the critical reviews of the problem by S. Steele from the Stanford Project (S 4.591, Footnote 5) and the skeptical views on word order problems in the Milwaukee conference of 1985 (UAJb 58.152-153). It is also self-evident that highly developed syntax (the first system in which the word order gains semantic importance) is evolutionarily the youngest part of grammar (this statement is actually already a self-contained universal), and therefore we can not expect that similarities in them are very old. Word order similarities came about, apparently, independently from each other (accidently, "processually") in different languages or language groups. The important elementary universals of language are certainly very old; we can even formulate this as a universal: the older a phenomenon, the higher the probability that it is universal. This leads to the conclusion that the evolutionary and diachronic perception of the language is, in my opinion, a methodological sine qua non of the research on language universals.

The list of the universals in this book contains virtually no cross-references. In some cases, I repeated universals which belong to more than one group (in most cases somewhat reformulated) rather than making cumbersome cross-references with numbers, titles, or topical definitions. The best way to label the single universals is with short and striking titles. Unfortunately, the title of the universals in my corpus are not always short – most labels (boldface print) run over one line. Nevertheless, any other system would have been even more disadvantageous. Numbering the universals in combination with the name of their discoverer (e.g. Greenberg 1, Greenberg 2, or Greenberg 18.7, done occasionally by some of the Stanford members) is certainly good if we want to honor a respected scholar;

however, numberings have little mnemonic value in a large data collection, especially for the uninitiated prospective reader and, hopefully, user of the present publication. Therefore, I refrained from numbering the single universals.

One of the difficult problems in the composition of this book was the systematization/classification of the language universals. The arrangement Presoundsequential –› Soundsequential universals follows the evolutionary pattern of how language developed from subhuman communication to human speech. The chapters on Phonetics/Phonology, Morphology, Syntax, Vocabulary, Semantics, and Olinguistcs run along the traditional line of grammarians (and, basically, of the Stanford publication). The short boldface titles on the right side in the text serve to orient the readers; they are not real chapter titles and offer topical qualifications mostly only with regard to the first universals and those following them. However, they may constitute the skeleton around which a more exhaustive topical classification of the language universals may be built. The casual chapter title Olinguistics (for the concept see GLC 85) may distress many serious linguists; however, I find its use as a comprehensive designation unavoidable for the numerous modern offspring of language studies like biolinguistics, ecolinguistics, ethnolinguistics, geolinguistics, macrolinguistics, metalinguistics, microlinguistics, neurolinguistics, pedolinguistics, preprotolinguistics, protolinguistics, psycholinguistics, and sociolinguistics, as well as zoosemiotics, pragmatics, and phyletic linguistics (phyletolinguistics). From these fields, of course, only a small collection of universals could be considered for the present publication.

I did not apply extreme terminological innovations (see what has been said above about the term universal/universe); nevertheless, I use such coinages as "uneme" for 'universal' (only occasionally) and "tocon" (from tone-icon); both are my creations. The term "soundsequential language" is almost the same as human language (i.e. "speech"); my view is that not the sound production but the soundsequence makes the human language. Sound production with signal relevance can be found in the animal world too; the arrangement of the different types of sounds in ordered sets (words) with clear semantics (and in large number) as a result of the so-called "concept pressure" is, however, only-human. I also refrained from the esoteric style so popular today in linguistic (also, specifically, in language universal) publications. Kenneth Pike emphasized in 1972 that "over-eagerness for theoretical contributions may... destroy the need for careful, extensive – perhaps even boring – gathering of data which, eventually, must be available for checking even the most inspired hunches concerning universals of language or of human nature" (Proceedings of the Eleventh International Congress of Linguists. Bologna-Florence, Aug. 28-Sept. 2, 1972. Edited by

Luigi Heilmann. Volume I. Bologna: Società editrice il Mulino 1972, p. 747; on p. 744, the same text ascribed, apparently by mistake, to Hiz). What the present publication offers, is data collection, nothing more. This happens also in accordance with the principles of the series in which this book appears. The publications in this series avoid, as far as possible, speculations and theoretical resolutions. It is the view of the author that the preference for theory building in modern linguistics since the 60s has brought about an attitude in which the description model, the interpretation, the method, the "how" became more important than the facts themselves and their systematic collection/evaluation. It is certainly time to return to data analysis and to disregard the copious extravagance of theoretical eagerness which disdains the real foundations of its subject matter: the simple linguistic fact.

The author addressed an audience in December 1967 at the University of California in Los Angeles, in September 1969 at the University of Helsinki/Finland, and then in January 1970 in Cologne/West Germany under the title "A Search for Universals in Finno-Ugric Languages" (see UAJb 42.9-17). In Los Angeles, the discussion called for a classification of language universals (H. Birnbaum), an important suggestion which remains a desideratum even today, after 20 years of research. In Helsinki, the late Valentin Kiparsky offered enhancing ideas from the point of view of historical linguistics. In Cologne H. Seiler (among others) participated with beneficial remarks in the discussion of the paper. After these events, the author remained for a long time at some remove from the enormous research efforts on language universals. However, he followed with interest the institutionalized investigations on the topics, writing reviews about important publications on language universals (so about the four volumes of the Stanford project and about Comrie's book, see UAJb 52.184-185 and UAJb 54.166). In the journal Ural-Altaische Jahrbücher-Ural-Altaic Yearbook (of which he is the editor), he printed from time to time reports on events of language universal research (see Anttila UAJb 42.268-271 and E. Moravcsik UAJb 58.152-153). Thus the author thinks that, after more than 20 years of observation and repeated review of the research results on language universals (which are extraordinarily important also for genetic, areal, and structural linguistics), the time is right for him to write this book and to put it on discussion in the rather large community of linguists who have ever dealt with the global aspects of communication structures. It was one of his purposes to include in the book a large number of universals which were found and described outside of the institutionalized projects on language universals – often overlooked by the representatives of the meritorious mainstream research. His view is that without good knowledge of language universals no real comparative studies (genetic, areal, or structural-typological) are possible.

At the end of this introduction I would like to remark that among the forebears of the language universal research we should mention the German scholar Adolf Bastian (1826-1905), a "staunch critic of Darwinism" who already at the end of the last century advanced the idea of "elementary ideas" (Elementargedanken) which may have been produced in the culture everywhere in the world by "general laws of the psychic unity of mankind" (see Sebeok below). Bastian's idea was taken up by the Austrian Hugo Schuchardt (1842-1927), who in 1912 introduced the notion "elementary relationship" (Elementare Verwandtschaft) in global oriented langu-age studies, pointing to what became the topical frame of modern research on language universals in the 60s and 70s. As far as I know, the ideas of Bastian and Schuchardt have never been appropriately acknowledged in the mainstream American language universal research (see Th. A. Sebeok in: Echanges et communications. Mélanges offerts a Claude Lévi-Strauss. The Hague: Mouton 1970, p. 614).

*

In my counting, the number of entries on pp. 21-136 of this book is 586; out of them, 88 concern presoundsequential, 204 phonetic/phonological, 116 morphological, 49 syntactic, 46 semantic, 36 vocabulary/lexemic, and 47 "olinguistic" language universals.

*

I gladly thank my wife Leena for her encouragement, understanding, inspiration, and technical help, without which this book could not have been written. For information related to a few special questions I am indebted to my colleagues R. Austerlitz and Th. A. Sebeok, and to the Budapest phonetician András O. Vértes. I also acknowledge the help I received in copyediting the text of the present publication from Evan P. Young of the Research Center for Language and Semiotic Studies. However, the responsiblity for the book lies exclusively with the author.

Bloomington, Indiana, October 23, 1987 Gyula Décsy

1. Presoundsequential Language Universals

Preliminaries: Prehuman/Subhuman

Speech (and, apparently, every Form of Communication) is Energy-Propulsion. We need energy for the production of sound (i.e. for pressing air from the lungs through the mouth). Compared with modern electronic (ectric) communication channels, the (human) language is very energy-intensive (i.e. inefficient). Energy-abundance, however, seems to be a common feature of biologically important functions of life (cf. procreation).

Sentientism. Difference in the information level between individuals is the driving power of information flow. Exchange of information (i.e. communication) is necessary/possible only between "sentient beings... who are different enough to consider each other aliens" (C. R. Peters & B. G. Blount in G 277). There is no need or chance for communication in a society in which the information level of the (extant or prospective) individual communicators is identical. A restless conveyance and total/complete reception of information would eliminate the need for any contacts whatsoever between members of a community. Theoretically, a total informedness as opposed to total uninformedness means, probably, a lack or cessation of life. Nevertheless, the trend toward or wish for uninterrupted increase in the information level of the communicators in a society is a basic feature of human nature. The sentientism (differing level of informedness) is certainly subhuman; signal flow exists already between genes in the chromosome and between individuals of the lowest organisms. The term "sentientism" is based on the observation that the "difference in information level" is a "feeling", an innate emotional concept of all organisms in the world. This feeling requires and causes signal flow (information flow).

In the Presoundsequential Time, Thought was Independent from Language. A certain primitive form of thinking (spatiality-oriented, without time/mood concepts or reasoning abilities) must be supposed for primitive species of the vertebrates. As these species have no (soundsequential) language, their "thinking" cannot be dependent on any form of language. A relationship (interaction) between thought and language can only be assumed for the soundsequential period of human evolution.

Phylogenetically and Ontogenetically, Thinking is Older than Language. This follows from the previous. The child already has thinking abilities before he or she begins to talk. The view that the world can not be understood without language or speech (extra linguisticam non est vita, si est vita non est ita) is simply not true. As Gipper emphasized, this monism is, from an evolutionary point of view, untenable/unwarran-table (K 547). The linguo-cognitive monism is based on the misinterpretation of a famous statement of Jean Jacques Rousseau (1712-1778):"la parole paraît avoir été fort nécessaire pour établir l'usage de la parole" (Decsy 1985.270). Today, in the time of computers and digital tabulations, we know that speech is not necessary for thinking. In most highly deductive disciplines, e.g. in mathematics, soundsequential communication has always been limited to a minimum.

Language does not Correspond Simply and Uniformly to Thought. "Even if there are conceptual universals, their existence is not sufficient cause for linguistic ones" (Brian F. Head in S 3.195).

It is Crucial to Keep Language and Speech Separate. Language is a deep structure phenomenon; speech is merely one of its possible surface manifestation(s) in the soundsequential (vocal, verbal) setting. Unfortunately, the average speaker (and even linguists) do not adequately distinguish the two concepts (namely language and speech), in spite of the much quoted suggestions of de Saussure (langage, langue/parole in French spelling; in German Sprache/Sprechen or Sprache/Rede; in the terms of American formalism competence/performance). The term "language" is used, very often, in the meaning "speech", and we cannot change this tradition. Language is the blueprint: speech is its vocalic-verbal realization by the speakers/communicators of a community. Language is the same as the score (the written form of a musical composition): speech is the same as the actual performance of the musical composition by an orchestra (Laziczius' metaphor). Language is the train scheduled to leave Geneva for Paris every day at 2:00 P.M. as listed in the timetable of the railroad company: speech is the actual train (it may consist of different railroad coaches every day) which leaves the Geneva railroad station (and need not be physically the same trainset every day). The latter is de Saussure's metaphor. The different meanings received many terms in linguistic research: here I give a small collection of those most commonly used:

parole	langue (de Saussure)	langage (French !)
Sprechen	Sprache	Sprechtätigkeit
Rede	Sprache	Redetätigkeit

Sprechakt	Sprache
äußere Sprachform	innere Sprachform (Humboldt)
Form	Inhalt (Weisgerber, Schlachter)
actual	potential (Jespersen)
Form	Bedeutung
speech	language
individual	social
solitary	collective
surface structure	deep structure
performance	competence
Realisation	System (German)
Sound	Phoneme
Formsystem	Semantik
practice	theory
etic	emic (Pike, cf. phonetic/phonemic)
variable	invariants (formalists)
definite	indefinite
dynamis	energeia (Aristotle Metaph. 5.1048)
concrete	general/generic
Wirklichkeit	Möglichkeit
Wort	Begriff
differentiated	generalized
empirical	speculative
greifbar	virtuell
inclusive	exclusive
implicit	explicit
one-time	universal

Speech has, in a certain Sense, a Phylo- and Ontogenetic Priority Over Language. Language as a blueprint of a communication system is an abstraction based on observations made of speech. The child conducts speech actvities before he or she acquires the so-called language competence. Modern linguistic schools (structuralism, formalism) stress the priority of language over speech. This may be correct from a consistent descriptive point of view; however, evolutionarily, speech has priority over language.

Language Emergence is a Problem Solving Event. "How to make explicit what is explicit in earlier abilities. Solving these problems results in the emergence of the symbolic [tone-iconic] world...by the invention of the conceptual mode" (A. J. Lock in G 1983.264).

Priority of Physical Contact

The Primordial Pattern of Communication is Physical Contact. It exists as a "given" relationship already in the anorganic world. Its basic form is the "fusedness": not the fusion itself as a procedure, but the motionless (timelessly existing) contact.

The Physical Contact of the Anorganic World is Condition and not Procedure. A condition as an enduring and immutable form of motionlessness does not require communication.

Communicatory and Linguistic Concepts are Traditionally Prescientific and Naive. They operate with "motionlessness" as a basic concept, although science has proven that there is no such thing as motionlessness (for details see Décsy 1981.2). We also state that "the sun rises" which is a pre-Copernician/Ptolemaic view. We have known since the 16th century that the sun does not rise but rather that the Earth rotates into the light beam of the sun in morning. Nevertheless, the metaphor "rising sun" (scientifically a nonsense) is used and will remain an indispensable part of every human language. Décsy 1981.2-3. See also "prevarication" below.

Communication as a Recognizable Procedure Exists only in the Organic World. The oldest organisms (bacterial fossils) came into existence on Earth 3.5 billion years ago; their appearance brought about the need for communication (Sebeok 1986.381).

Proto-Touch. Physical contact as the oldest form of communication can be termed "Proto-Touch", "Ur-Touch", or (at Révész as a German coinage) "Urkontakt" (GLC 123-124). In the early developmental stages of the organic world the "Proto-Touch" or its lack already may have contained the basic semiotic values good/bad, like/hate, caress/beat.

Physical Contact as the Primordial Pattern of Communication is Even Today Irreplaceable in Basic Biological Procedures. E.g. in procreation (GLC 123); in the acceptance of nutrition; in breathing; in highly emotional relationships like huddling, petting, caressing, licking (already in the animal world) and kissing, handshaking, slaps on the shoulders, etc. (in humans). A phone conversation with a person to whom we feel attracted never gives the same satisfaction as meeting in person and bodily touch (embracing, kissing, handshaking).

Most Synaesthetic Metaphors are Tactile-Based. Ullmann UL 243. This also proves the high evolutionary priority of touch in communication.

Non-Tactile Sensical Contacts

Every Type of Communication Imitates Physical Contact. Every form of communication is only a replacement/surrogate for the missing and desired physical contact. There is an innate drive in every organism to communicate in the original sense of this term, i.e. to establish physical contact (see the concept "Urtouch"/Prototouch" in GLC 123).

Every Type of Communication Other than Physical Contact can be Regarded as a Deception/Illusion because they pretend to establish a physical contact without actually doing so. The establishment of real physical contact would make the alternate type of communication unnecessary. Physical contact (touch, "tacteme") is the most perfect form of junction (fusion) between two communicators.

Traditionally, only Non-Physical Contacts are Regarded as Communication. They can be grouped along the evolutionary line of the senses: taste, smell (olfactory), hearing, and vision – all subhuman (and, partly, submammal) abilities.

E(le)ctric and Anorganic; Gravity

Exosensic Ways of Message Conveyance (Ectric Communication) played and play a Major Role. These are the various ways of electric and electronic communication conducted (as a rule, between robotic communicators) beyond the sensual reception abilities (touch, taste, smell, hearing, sight) of humans. They exist already in the anorganic word; however, they remained outside of the conscious sensual communicatory systems of man except for such phenomena as lightning or (electro)magnetism. The suprasensual ectric communication of our modern world is a return to the oldest (prehuman, and probably preorganic) forms of message conveyance which are simpler and more effective than any (real-tactile or pretended) contact through the sensually based channels (visual, auditory, olfactory).

Gravity Played a Major Role in Language Development. The upright posture of man is understandable only in the relationship of man to the Earth (gravity). Results of upright posture (and of gravity) are: the specific human shape of the face and brain, the F-like arrangement of the pulmonary-laryngeal-oral tube of sound production, separation of the nasal cavity from the oral area (velopharyngeal closure – Décsy 1985.267), closeness of the upper and lower lips to each other (and to the teeth), reshaping of the jaw, shortening of the tongue, diminishing the atmospheric pressure on the human body and especially on the mouth (which eases expiration compared with the breathing of quadrupeds). All these features are human-species-specific (only-human).

Electromagnetism Plays a Major Role in Speech Production. Out of the four fundamental (anorganic) forces of nature (gravity, electromagnetism, the strong force which holds the centers of atoms together, and the weak force, which makes some atoms break down in radioactive decay), electromagnetism plays a major role in the functioning of language and the production of speech. Electromagnetism is responsible (together with many different chemical procedures – see below) for communication of the brain cells within the nervous system and for information storage there. On newer (less important) forces (fifth and sixth forces) of nature see the Louisville (Kentucky) Courier Journal (Dec 12, 1987, A3) reporting on announcements concerning discoveries in a meeting of the American Geophysical Union by Donald Eckhardt. An increase of electromagnetic activity in connection with reasoning may be a human-species-specific feature (animals have fewer "blips" than humans – see GLC 81). Electromagnetism seems to be responsible more for "rational" than for emotional bits.

Chemistry: Taste/Smell

Chemical Procedures are of Crucial Importance in Brain Activity. Besides electromagnetism, chemistry plays a major role in cognitive activities and speech production in the human brain. These chemical procedures are wired into the (archaic) limbic system (GLC 81) and serve/cause mainly emotional activities. Brain cells communicate by "drips" with each other; their lack or abundance affects behavior (drug-related problems). The chemical balance of the body can be influenced by varying the intake of drugs.

Taste is a Differentiated/Specific Variation of Touch. It still requires (like touch) a direct physical contact (food on the tongue in the mouth cavity). From a semiotic point of view, taste ("tasteme") can be regarded as a chemical sign (one of the "pheromemes", – for the term see Décsy 1981.27-29).

Smell (Olfactory) is Evolutionarily the Oldest Signal in Subhuman and Human Communication which does not Require a Direct

Physical Contact. However, it still signals bodily closeness – in the present or in the not too distant past. Smell is a rapidly fading (evanescent) sign; thus it displays one of the typical characteristics of the Hockettian list of human language features (see below). Smell (as "olfactoreme", "odoreme") is also a chemical sign (pheromeme).

Audio/Hearing

Audio (Hearing) is, After Smell, the Second Oldest Form of Distant Sign Reception. Its reception organ is the ear, extant already in the vertebrates and (in rudimentary form), in some sound producing insects. It is a highly specialized part of the body surface with complicated neural organization. "Earshot" (auditory range) is, generally, more extensive than the perception area of smell. Based on the wave diffusion principle, the ear accepts indiscriminately any noise from an area within a certain distance (earshot). A noise functioning as a sign can be termed "audieme" (Decsy 1981.29-32). Normal humans accept audio signals between 16 Hz (hertz) and 20,000 Hz (speech between 150 and 10,000 Hz, music between 50 and 20,000 Hz). The (very old) hearing ability of humans challanged them to produce noises. This may have been one of the stimulating factors in the startup (development) of human speech (noise, then sound) making.

Vision/Sight

Vision Is a Significant Factor of Human Communication. Even if not directly involved in the vocal-auditory channel, vision is an important (today, probably the most important) source of information. If not alphabetically written (like this text), visual signs ("visemes") do not require a soundsequential encoding. Visemes are iconic: pictures, traffic signs, and the entire world we see around us in its physical reality or in pictures. Visemes represent the most direct communicative experience. Today, about 80 percent of the information a normal average human receives is viseme-based ("sometimes a picture is worth of thousand words"). In the most developed form of human vision, the color spectrum, the frequency borders are between about 380 nanometers (violet light) and about 700 nanometers (the boundary where red light grades into infra-red radiation – for details see GLC 124). See also the important universal: sound is gesture (below). As far as sound is a gesture, it *is* recognizable (decodeable) through the visual channel by trained persons. (Hearing impaired people identify sounds on the basis of the movement of the lips and mouth muscles of their interlocutors.)

Reading is not a Form of Visual Communication – at least in languages which use soundsequentially-based alphabets (as most European languages do). These alphabets are soundsequentially coded, and time sequence plays a role in their decoding. Chinese and Japanese text reading is in principle soundsequence free, and therefore (from a technical point of view) superior to European alphabetic reading. The brain processes iconic scripts (pictograms) in the right (non-verbal) hemisphere without the participation of the Broca's and Wernicke's areas.

In Visual Communication, Humans Signal by Motion as in gestures, mimes (mimics). Sign production of this type cannot be carried out in a condition of motionlessness. However, motionless pictures may have signal value too; however, the signal value of motionless pictures is more limited than the utility of motional semiosis.

Gestures

Gestural Affirmation is Universal. Affirmation is done by nodding (lowering and subsequently raising the head back to the starting position on a virtual horizontal axis at the neck-height between the shoulders) as a sign of agreement or acknowledgment. Known exception: Bulgaria/Turkey and some neighboring areas, where the affirmation is done by a slight motion of the head to the right on a virtual horizontal axis between the front and the back part of the neck. The Bulgarian affirmative head move is often understood as a negation by uninitiated foreigners in Bulgaria.

Gestural Negation is Universal. It is done by a single left-right motion (mostly left first, right second) of the head on a virtual vertical axis in the center of the neck. Exception: Bulgaria/Turkey and some neighboring areas, where negation is expressed by an abrupt raising of the head back on a virtual horizontal axis at the height of the neck between the shoulders. The Bulgarian/Turkish type negation is very similar (though never completely identical) to the universal affirmative gesture. Nevertheless, it is often misunderstood as affirmation by uninitiated foreigners in Bulgaria.

Shaking the Head in Disapproval is Universal. This is a series of left-right motions of the head around a virtual vertical axis in the center of the neck. It differs from the gestural negation that not just one motion, but a set of motions (four-five) is performed in an uninterrupted sequence. This is certainly a variation of the negation. However, the repetition of the motions (quite surprisingly) decreases (and not increases) the intensity of the negation (disapproval).

Gestural and Vocal Channels in Human Communication have the Same Underlying Processes. The gestural origin of language/speech is an old thesis of linguists, most recently represented mainly by Wescott, Hewes, Stokoe, and Harnad (see Décsy 1977.82). It is regarded today as proven (see also lateralization displayed by right-handedness and by Broca's -Wernicke's area). I quote here the formulation of A. J. Lock: "The two phases of [cognitive] development, the gestural and the linguistic [vocalizative] are firstly being seen as essentially continuous, and secondly, in consequence, as based on the same underlying processes" (G 261).

Signal Noise Production

Noise Production is Common in the Animal World. It is used for signaling (semiosis) too. Some noise producing insects can "hear" (even without specific "ears", as e.g. the crickets). With this, the existence of a rudimentary subhuman pattern for the Vocal-Auditory Channel (see below) is proven. Note, however, that the noise-producing apparatus at this subhuman stage of the development is often not the lung/mouth tube, but rubbing together external body parts (male crickets) or drumming with the "hands" (gorillas).

Vocal-Auditory Channel, the main means of intraspecific soundsequential human communication, is a phylogenetically old phenomenon. It is most developed, however, in bipeds (birds, humans) or in animals which are able to be (at least halfway) in upright body position (penguins, some frog types).

The Hearing Section of the Vocal-Auditory Channel is Much Older than its Vocal Section. Humans (and subhumans) were already able to hear long before their sound (noise) production apparatus was developed.The hearing apparatus of a newborn human is completely developed (the visual sense of infants reaches completion only in the third month after birth). Phylogenetically, this hearing ability may have stimulated the development of the sound (noise) production apparatus.

"Sounds are Fitted, as Colors are not, for Use in Signaling Rapidly, subtly and in an infinite variety of ways. An animal can pass in an instant from silence to sound, from high pitch to low, from harsh noise to pure sweet tones, from loud notes to soft ones; and it can produce such changes in all sorts of different sequences and patterns" (Hartshorne 1973.14).

Singing of Birds may have been a First Class Stimulating Factor in in the Emergence of Human Vocalization. Birds have a much greater phylogenetic age than primates (and especially humans). Their impressive singing is the result of a much longer period of development than is the noise (sound) production of human articulatory organs. The singing of birds is technically more perfect and self-contained than human vocalization. This is mainly due to the syrinx and a special airbag which makes bird-singing on a large scale independent from breathing.

Sound Production is an Exaptation (Secondary Activity) of the Articulatory Organs. The primary function of lungs, larynx, pharynx, nose, mouth, tongue, teeth, and lips is breathing and nutrition intake. Sound articulation superseded these primary functions of the organs in question, but it left naturally intact their original functions. There are no organs in the human body which would serve only for (secondary) noise production. All articulatory organs in humans kept their primary biological functions. For the concept "exaptation" (a remodeling of the term "adaptation", see Sebeok 1986.169).

As a Resonator and Articulatory Factor, the Larynx-Pharynx-Tube is Older than the Mouth Cavity. The resonant variability of the larynx/pharynx (subvelaric) area in some subhuman mammals is more developed than in humans. Subhuman sounds are produced primarily in the pharynx/larynx tube and not in the mouth cavity. The mouth cavity as a variable main resonator is only-human. Nevertheless, laryngeals and pharyngeals are still used in present-day human languages too (in a rather differentiated manner, e.g., in the Semitic languages). Subvelar (laryngeal/pharyngeal) sound production techniques are highly developed also in the Khmer and Mon-Khmer languages in Southeast Asia (see Gregerson 1986).

The Applicability of the Mouth Cavity as Resonator in Sound Production is the Result of Bipedality. Upright posture made the shift from the more archaic (basic) laryngeal sound production to oral sound production possible. Animals are unable to use the mouth cavity as a resonator in signal noise production. Animals are habitual nose-breathers: only humans can breathe both through the nose and (as an innovation) through the mouth, or exclusively through the mouth (which is a precondition of oral sound production – Décsy 1977.28-29).

Airbaglessness is a Precondition of Human Language Emergence. Birds and also some primates have airbags, and use them in the same way that humans use the lungs in connection with sound production. Such animals talk almost uninterruptedly because their sound (noise) production is not dependent on breathing. They do not have to stop speaking/singing and thus have no time to contemplate and meditate before they "say" something. The result is that their talk has little or no communicative function. The ability of humans to think is due to the dependence of talking on breathing (using the lungs in an unnatural way as an airbag); this compels humans to make pauses in their speech and leaves them time for consideration and contemplation.

Exhalation/Expiration is Less Energy Consuming than Inhalation, at least in the upright position of the human body (or of its upper part). That is why sound production is expirative, and why we speak most conveniently when we stand or sit. For the physical background of this see Décsy 1977.29.

Laughing, Crying, Snoring, Coughing, and Sneezing (Sternutation), and Kissing are Universals. They have the same semantic value among humans everywhere in the world. They may have stimulated the introduction and development of pulmonary sound production in humans. Most of these sound products are expiratory (pulmonary), except the snoring (which can be inspiratory, but is, as a rule, expiratory).

Pulmonary Vocalization

The First Communication from the Infant to the Adult is the Cry. Reich 1986.33. Phonetically, the cry is an unarticulated pulmonary laryngeal (vowel/conconant) of a long duration. Its production technique may have been one of the factors stimulating humans to consider similarly pronounced "sounds" for communication.

Human Sound Production is Monogenetic. It is conducted in the same way by all normal (healthy with regard to speech organs) human beings of the world. The "single sound" production abilities of humans were developed very early (probably several hundred thousand years ago, perhaps in pre-Neanderthal times) in the oldest undivided clans somewhere in Western Africa which are commonly regarded as the ancestral units of the entire human race. It is possible that Neanderthal man already had timbric abilities; he may have been able to produce vowels like u, i, and a as well as to pronounce some distinct laryngeal/pharyngeal consonants; perhaps also some dentals, but no distinct labials except for w (his lips were still too far from each other and from the teeth).

Coarse Vocalization is the primary form of pulmonary signal noise production; it lacks "in delicacy or refinement" and uses mainly stress, tone, and length variations without timbred vowels (H. Jakob and W. K. Müller in G 1983.250). This is the typical sound production of newborn humans, and was the primevial "speech" of mankind before the Cro-Magnon introduced fine-modulative articulation with the timbred vowels /a, u, i/ and soundsequential morphemics. Coarse vocalization is common in the animal world; its phonic realization is, in most mammals, an inarticulate noise (sound) which makes an acoustic impression (and has a frequency) similar to that of human laryngeal "vocsonant" /ə/ʔ/. Occasionally, especially in birds, the sound product can be (with regard to the acoustic impression) similar to the timbred vowels /i, u/.

Tone/Pitch is Subhuman. It is the most important vehicle of vocalization. It appears in the early monosyllabic babbling and cooing of infants "which develops into complex multisyllabic utterance, complete with intonation pattern" (Reich 1986.33). Neurologically, tone is centered in the cerebellum (an archaic part of the brain) and not in the speech center (which is located in the evolutionarily new neocortex). No human noise product is, therefore, emotionally so strongly saturated as tone.

Stress is Subhuman. No sound, syllable, or word can be produced without stress (the force with which a sound or syllable is spoken). In this sense, every syllable, even the "unstressed", has a stress. The "unstressed" syllables merely have less (weaker) stress than the stressed syllables. "Stressed" and "unstressed" constitute thus gradual and not equipollent oppositions. Stress also has its neurological center in the cerebellum.

Quantity Opposition is an Archaic Feature in Vocalization. There are two ways it comes into being. **1**. A sound of long duration is interrupted: iiiiiii ---> i-i-i-i-i-iii (many further combinations are possible). **2**. Short sounds are repeated. Concerning brain topology, quantity opposition is subcortical (like stress and tone).

Concept and Concept Pressure

Concept is Older than the Sign. It is logical that we cannot name a thing (concept) which we do not yet have. We are able to dispatch a sign only to a concept which already exists (as a result of experience) in our mind. As thinking is nothing but (re)cognition of relationships between a large number of concepts, the high priority of concept over sign (and overlanguage as a system of signs) must be regarded as self-evident.

Concept Pressure. Concept pressure (Begriffsdruck) is an accumulation of a large number of new concepts which do not yet have names. The large number of concepts compels humans to invent new names (Décsy 1981.10-12). Large number of words (soundsequences with semantics, vocabulary) came into being at the time of speech origin as a result of concept pressure (theory of inventionalism, represented by F. R. H Englefield – see Décsy 1985.269).

Concept Pressure Leads to Invention/Acquisition of New Signs. This is, apparently, already an intellectual activity. If I know a thing, I must have a sign (name) for it. Invention under concept pressure is one of the important factors in early (and modern) sign production. This was recognized by the British linguist F. R. H. Englefield (1891-1975), see Décsy 1985.269. Concept pressure is observable in infants in the first months after birth.

Sign Emergency (Zeichennotstand) is a Permanent Feature in Human Language Communities. Every expanding human society has more (new, innovative) concepts than it has names (words). From the beginning, humans have always had problems (sometimes more, sometimes less) with naming new concepts which were produced as a result of social and economic development. Almost legendary are the strange methods of place-name creation in the United States in the 19th century, when there was an enormous need for distinct designations of newly founded settlements (and there is even today such a need, cf. "Sound-Alike Streets Put Suburbs on a Road of Confusion." Chicago Tribune. Section 2 – Chicagoland, Sunday, November 22, 1987. pp. 1-2).

Concept and Sign Types

Two Basic Forms of Connection Between Sound and Sign. Such connections are, from an evolutionary and logical point of view, either causal or arbitrary. The additional sign types (subtypes) of Peirce and Sebeok (signal, symptom, index, symbol, name, and emblem – see GLC 100), are only subgroups, either of the causal signs or of the arbitrary signs. In the terminology applied in this book, icon is identical to causal sign, and symbol is identical to arbitrary sign. Here I give a chart of the terms used in earlier research for the denotation of the arbitrary/causal connection between sign and concept (after Décsy 1981.16):

physei	thesei/nomoi	Plato
true	false	Greeks

realistic	formal	Gombocz
natural	artificial	Wundt
Urschöpfung	(Neuschöpfung ?)	Hermann Paul
true	arbitrary	Plato
correct	(conventional ?)	Plato
causal	arbitrary	structuralists
motivated	unmotivated	R. Anttila
concrete	abstract	semanticists
descriptive		
expressive		
icon	symbol	Peirce
Lautmetapher		Wundt/H. Paul/Fónagy
soundsymbol (!)	(?)	Jakobson and followers
tone-icon	(tone-symbol ?)	Wescott
tocon(ic) (< tone icon)	arbitrary	
tocon word	arbitrary word	Décsy
natural	conventional	

In Human Language, Arbitrary Signs Prevail. Every causal word becomes an arbitrary word after a certain period of use (see *crush* in English or *Krieg* in German). After conventionalization, the causal character (natural-iconic origin) of every (originally non-arbitrary) word is forgotten ("de-iconization", "decausalization" – GLC 48). This procedure does not take very long.

In Pre-Human Communication Systems, Arbitrary Signs Prevail. Except for the dance of bees, no causal signs have been considered for the animal word (Décsy 1981.21-22).

Arbitrary Signs Precede Iconic Signs. A causal sign (icon) cannot be understood without highly developed intellectual (non-rudimentary) reasoning. As reasoning is a late achievement of humans, the iconic sign cannot be primary (Décsy 1981.21-22). The arbitrary sign must be older than the iconic sign.

Primitive Arbitrariness is Identical with Emotionality. As primitive sign recognition could not be reasoning (causality) based, it must have been emotionally/instinctively based. This is the area of human development in which innate/inherited/instinctive/unexplainable/mythic elements play a major role.

Intellect is Controlledly De-emotionalized Arbitrariness in Production and Understanding of Signs. This is, probably, the feature which makes humans.

Every Single Human Language Contains Universals. "Suppose that all the languages of the world except English were to become extinct. Thereafter, avery assertion true of English would also assert a (synchronic) language universal" (Hockett UL 3). In principle, language universals could be identified only on the basis of investigations of a single language.

Sign System Types

Martian/Venusian/Tellurian Sign Systems and Human Language. The most primordial relationship between concept and sign was "Martian": therein every distinct concept corresponds to one and only one sign (=proto amorphic type, knows only full-lexemes). "Venusian" is an imaginary system, in which every concept could be expressed by every sign (=total chaos). The human (=Tellurian) language is in between the two extremes. Subhuman communications are basically "Martian", as are the modern computer operations (note: most developed systems return to the most primitive ones !). The terms Martian/Venusian/Tellurian after John Algeo (1970) and A. Makkai (see S 3. 405-406).

Brain as Speech Organ

The Speech Center is Unilaterally Positioned in the Brain. It is, as a rule (in 90 % of the world population), in the left hemisphere of the brain (neocortex, neopallium, called Broca's and Wernicke's area) in all healthy and normal humans. It is in the right side only in humans who have their heart in the right side of their chest. Left-handed people have their speech center as a rule also in the right (and not in the left !) brain hemisphere.

Broca's (Motor Speech) Area is the Center for Speech Production. It is placed in the neocortex in the posterior portion of the left frontal lobe. Described in1861 by Paul Broca (1824-1880), this evolutionarily young area is responsible, apparently, mainly for the production of timbric vowels and non-laryngeal consonants. Cf. C. Romero-Sierra: Neuroanatomy. A Conceptual Approach. London, etc.: Churchill Livingstone, p. 360.

Wernicke's (Sensory Language) Area is the Center for Sound-sequence Comprehension/Recognition. "This cortical area is composed principally of the posterior one-third of the superior temporal gyrus".

Described in 1876 by Carl Wernicke (1848-1905). Romero-Sierra, op. cit. p. 360. Sound-sequence-based information received through the Vocal-Auditory Channel are processed in Wernicke's area.

Cognitional Duties of the Bicameral Units. A rigid bicameral division of cognition procedures is rejected today. However, it seems that there are certain functional differences in cognition procedure between the two hemispheres. The left is the center for denotative or serially ordered phonological-grammatical-referential processes; the right... is decisively involved in connotational (gestaltlike, orchestrating) complex processes ruling the combinative use of left-side units (Armstrong-Katz in G 1983.229). Left/right are responsible mainly for the following:

left	right
sounds/phonemes	affects, emotions
morphemes	elaborated connotations
restrictions	synthesis
names	logical analysis
entire words, word parts	metaphors
soundsequential codes	myths
rationalized messages	syntax, combinations
de-emotionalized information	music, poetics
paradigmatic sets	individual-creative activities
analysis	comprehensive evaluations
soundsequence writing	Chinese and Japanese script
	ideograms, icons, pictures

The quasi-commanding position of the right side was stressed in 1976 by J. A. Paredes and M. J. Hepburn (see details by D. Armstrong and S. H. Katz in G 1983.231-233). Probably, a large amount of non-soundsequential information is received and processed by the right and left hemispheres outside of Broca's and Wernicke's area. All this means that the soundsequential speech center of the brain is a recent development, and does not play the absolute commanding role in communication, as the adherents of the theses on the uniqueness and great age of soundsequential language assumed.

The Transmission of Information (Stimuli) through the Ears to the Brain is Contralaterally Organized. Soundsequences (morphemes, words, etc. – see left side list above) reach us mainly through the right ear; music, verse, etc. (see right side list above) penetrate into our brain mainly through the left ear. However, there is no consistent separation/

unilateralization with regard to the two information groups, and individual differences are considerable.

The Speech Center and its Unilateral Positioning in the Brain is Human-Species-Specific. No such unilateralization could be proven for subhumans. For references see Décsy 1985.268.

Right-Handedness and Unilateral Positioning of the Speech Center in the Left Hemisphere of the Brain represent a Contralateral Connection. Right-handedness may have influenced the positioning of the speech center in the left hemisphere of the neocortex, perhaps because speech was originally used mainly in connection with activities conducted by the right hand.

There is no Speech without a Speech Center in the Brain. The speech center (Broca's-Wernicke's area) is the most important articulatory (and speech) organ. Animals (especially higher primates) may have similarly arranged "articulatory" organs (lungs, larynx-pharynx, mouth cavity, tongue, teeth, lips); however, as they lack a developed speech center, they cannot learn to produce (articulate) sounds, words, and sentences. Speech production has its mental preconditions. Animals (even the higher primates) lack these mental preconditions; they are unable to develop speech centers (represented in the human brain by Broca's-Wernicke's area) or nervous systems which can move tongue, lips, jaw, larynx/pharynx and lungs in a "fine-tuned" way as humans do to produce articulate speech. The mouth-pharynx-larynx of animals are not "exapted" for speech; their exclusive natural functions are (as in earlier humans): food intake and breathing.

Tone, Quantity, Stress and Laryngeals are Subcortically Wired in the Brain. They are centered not in the neocortex (as timbric vowels and most consonants), but in the cerebellum (limbic system). That means that tone, quantity, stress, and laryngeals are older than the human (primate) neocortex, and in particular, older than Broca's and Wernicke's area. Speech impaired humans who are unable to speak an articulate language can still produce laryngeal vowels/consonants, and even vary them according to stress, quantity, and tone (pitch).

In Human Vocal-Auditory Communication, there is a Connection Between the Speech Center and the Subcortically Wired Sounds. If the laryngeals and the suprasegmentals (tone, quantity, stress) are applied in language, they are subject to the command (by nerve pathways) of the speech center. This is not the case in non-linguistic (unarticulated) sound production (sudden cry, laughter, etc.).

Bit Storage in the Human Brain

Bit Capacity of the Human Brain. According to estimates, the human brain (which weighs an average of about. 1,6 kg) is able to store $5x10^{10}$ or $2{,}8x10^{20}$ bits (Décsy 1985.25).

Two Types of Associative Connections and their Role in Bit Storage in the Brain. The semantic/iconic concepts are certainly primary (wood helps to recall tree; the daisy is associated with carnation, because both are flowers; chair has something to do with table, because both are parts of our kitchen furniture); grammatical (morphemic) association (the -er in thinker and keeper which connects "thinker" and "keeper"; or think in thinker and thinks which connets these two concepts) are more recent developments. If we recall wood, tree comes in mind with it; if we pronounce the word thinker, both keeper and think are somehow within our mind because of the morphemic associative connections (see Ullmann UL 222). According to more recent views, morphemes or even larger units (concept images) are stored in the brain, and not phonemes or their parts (as analyzed, e.g. by the adherents of the distinctive feature theory).

Concepts (Semantic Units) and not Soundsequences are Stored. The soundsequence is a coded form of the concept; it is not a basal form of storage in the brain. Soundsequential (or, more probably: morphemic) storage is possible in languages with highly reading and letter oriented memory systems (like most Europoean languages). However, the natural memory bit is concept oriented, amorphic, and "analphabetic". Broca's and Wernicke's areas are not storage rooms; they are only nerve centers responsible for en- and decoding of soundsequences (Broca's area for soundseqence production/ encoding, Wernicke's area for soundsequence recognition/decoding). Therefore:

Soundsequence Similarities are never Disturbing for Native Speakers. Native speakers never realize soundsequence similarities (homophonous forms) of different concepts (e.g. English pan vs. pain, Finnish kissa 'cat' vs. kisa 'sport game', Hungarian kap 'he/she/it receives' vs. pap 'priest') because in their minds the concept and not the (soundsequential) name is stored, and the visual impressions of these concepts are not similar. However, native speakers easily confuse things which are soundsequentially (in tactile, olfactory, visual, auditory) experience not similar, even if the particular concepts have distinct (soundsequential) names.

Soundsequence Similarities are Disturbing only for Second Language Speakers. This, however, is true only if they learn the language from (soundsequentially based) books; their primary memory imprint remains then for a certain time soundsequentially coded. The soundsequence codings are kept, apparently, in the short term memory storage rooms of the brain. It always takes a long time (and experiences with the real world in the new language community) before concepts coded only soundsequentially are transmitted into the permanent memory rooms. It is remarkable that puns are made (and recognized as effective word games) mainly by freshly assimilated foreigners.

Permanent Memory is always Semantically Coded (Lemme and Daves in Lass 1.368-369). This points to the high evolutionary priority of semantics over soundsequence.

Soundsequential Encoding is Inefficient. A memory bit can contain a sound, a visual concept (e.g. a house), or an entire paragraph of a speech or book. Although the memory bit is unexpandable, its storage capacity is enormously different (variable) according to how and what kind of concept is stored in it. It is clear that if we use up the bits for such meaning-inefficient elements as sounds, we do not proceed economically: for /opera house/, e.g. we need nine bits if we use a soundsequential encoding but only one bit for a visual-iconic encoding. The sensory-motor connection with this visual-iconic bit needs only a single reflex, and not a long decoding procedure running through a soundsequence in which time (sequencing the sounds after each other in a certain order) is always involved. Semantic memory bits do not need the time factor for en- and decoding.

Prevalence of Soundsequence-Oriented Storing is the Rule Only In Second Language Speakers. For second language speakers, there is generelly a direct (or only possible) connection between soundsequence and concept – if the foreign language is acquired through the method we use in our schools (priority of soundsequence and grammar). At the first stage of foreign language knowledge, there is no direct connection between concept and language use – the encoding is not reflex-like in the vocal application. Later on, the obligatory de- and encoding procedure can become the rule, and the second language practice will be similar to the first language use. Soundsequence-oriented coding is typical for the short-term memory process (see Lemme and Daves in Lass 1.368).

Language Acquisition and Language Attrition

First Language Acquisition is a Systematic Build-Up of Broca's-Wernicke's Area in the Brain of the Child. First language acquisition is the way infants learn the language from their parents or caregivers. The procedure is absolutely universal in all language communities in the world (Reich 1986).

Language Acquisition and Language Attrition are Reverse Procedures. This is a phylogenetically and ontogenetically valid universal, both for individuals and for language communities. Newborn infants learn first to use "laryngeal-velar consonants/vowels" strongly intervowen with stress, tone and quantity (longer or shorter singings/cryings). After the sixth month, the vowels i/u/a, (w-like) labials, and nasals follow, then the dentals, and at the and the o/e and the complicated consonants (fricatives, dentals, and finally k, which they spoke in the subhuman phase of laryngealization but later forgot to pronounce and have to relearn). Then grammar and syntax follows. In aphasia, humans forget the language in the reverse sequence: first the more complicated syntax, then some morphological forms, then the affricates, and so forth. At the end, only the laryngeals remain in the form of (reflexive) crying/laughing. The principle of this universal was recognized by Roman Jakobson in the 1940s. See his classic work in German: Kindersprache, Aphasie und allgemeine Lautgesetze. Uppsala Universitets Arsskrift 1942:9.1-83.

The First Sub-Human Sound Product of the Infant is the Ideomorph Consisting of Laryngeal Vowels and Laryngeal Consonants. This has been stated recently by Reich 1986.37. The sound complex can be rendered by the phonetic formula /ʔəhəʔ/ or /ʔəʔ/ I find this important because it confirms my suggestions from 1977 concerning the (undivided) laryngeal (vocsonant) as the phylo- and ontogenetically primary sound product of the human language (Décsy 1977.56). I knew families in Finland who decided to nickname their children with the first sound(sequence) which they (the infants) produced. One of the children is called Oko, and the other is called Nökö (in latter, the initial n seems to be a rendering of velar nasal which is close to the laryngeal consonants).

The First Human Vowel Products of the Child are a i, u. Stated after Reich1986.55. Different children acquire the three vowels, apparently, in different sequence.

The Earliest Consonants a Child Masters are p, m, n, h, w and b. This happens latest at the age of 18 months (Reich 1986.57. The supposed phylogenetic sequence here is h, n, m, p/b.

Children Generally Acquire I Before You. E. Clark S 1.255.

Societies Lose their Language in the same Way as Individuals. Speakers first become insecure with regard to the most complicated syntactic forms (learned last). At the very end, only the characteristic articulatory features of the (lost) first language are kept (intonation, quantity correlation, stress system). This is a self-evident universal: societies consist of individuals, and the language attrition of a society is a subsumption of the language loss of its individual members.

In Baby-Talk the 1st and 2nd Persons are Often Replaced by the 3rd Person. Ferguson S 1.213, after him Brian F. Head S 3.169. (Head uses "nouns" instead of Ferguson's 3rd person which is, basically, the same). This is certainly an ontogenetic universal; phylogenetically, however, it may be secondary. The 3rd person belongs to the human phase of language development. When the baby makes a difference between three persons and nouns, then he/she is already in the human (and not in the prehuman) phase of language acquisition.

Recapitulationism

In Language Development, Ontogenesis Recapitulates Phylogenesis. This is one of the most important results of recent Language Origins Research: the human language went through the same developmental stages during its entire history as the newborn human goes through between birth and acquisition of complete linguistic competence (approximately in the 6th year of life).

The Phases of Phylo- and Ontogenetic Parallelism are Different (Disproportionate). The subhuman phase was phylogenetically very long (perhaps several hundred thousand years), but it is short (about six months) ontogenetically. The human phase is phylogenetically very short (only about 30 thousand years); however, it is very long ontogenetically (the entire life of a human being except for the first three to six months after birth). This recognition was made possible by the achievements of modern pedolinguistics (child language research) during the last decade. Pedolinguistics (see the excellent handbook Reich 1986) became one of the most important auxiliaries of Language Origins Research.

2. Soundsequential Language Universals

Generic Tenets

Not the Sound but the Soundsequence makes the Human Language. Sound production through the lung-mouth tube is very common in animals (especially birds): soundsequence production is, however, only-human. Human language begins with the soundsequence and not with the sound itself.

Soundsequence Production is Certainly Polygenetic. Sound-sequences (words) were probably set up independently in the many clans and small human communities of the fourth-third decamillenium. The differences between the many thousands of languages of the world consist mainly in the sound sequences (words, grammatical forms), and not in the articulation of the sounds.

Thought (Cognition) is Independent from Language in Soundsequential Times as well. Valid only if we understand language as "speech", i.e. as a soundsequential system based on en- and decoding and running through the vocal-auditory channel (for this see Hockett in HC 1982.7, diagram 1); rudimentary thinking with non-emotional elements is well known in the animal kingdom in more developed species which have no soundsequential language. The favorite discussions of French, German, and Russian philosophers and linguists since Rousseau (filling many yards of bookshelves in our libraries and dogmatizing the view that "extra linguam non est vita") were probably the most futile attempts of modern scholarship. Thinking is not bound to language knowledge (see my remarks in Reviews in Anthropology 12[1985].270). In Germany, thought is even today often identified with language (especially by Lorenz and Hopp); cf., however, the more skeptical view of Helmut Gipper, who speaks about the "unresolved problem between language and thought" (K 548) in his excellent contribution "Evolutive Grammar or How to Become a Competent Speaker of a Language" (K 541-548).

Every Single Human Language Contains Universals. "Suppose that all the languages of the world except English were to become extinct. Thereafter, every assertion true of English would also assert a (synchronic) language universal" (Hockett UL 3). In principle, language universals could be identified only on the basis of investigations of a single language.

Duality of Patterning is Possible only in Soundsequence. In the soundsequence, the sounds have much higher signal (significatory) value than in separate use. They maintain their phonetic value (i, u, a, p, t, etc.) and they receive (as an addition) a positional value: the difference between meat and team (a duality of patterning, see Hockett's features below) is possible and/or relevant only in soundsequence.

Soundsequentiality Stimulates the Extension of the Phoneme Inventory. Once the usefulness of the sounds in the sequence had been recognized, people began to invent and introduce new phonemes in order to increase the number of possible combinations and create new names for things (concepts).

Only a Systematic Soundsequential Language can be Regarded as a Real Natural Human Language. Sound-like noises produced in the lung-mouth channel are known in small number in most mammals; they are often used for communication (dolphins, birds), representing clear-cut signal structures. However, their semantic pattern is limited. Only humans are able to build up a system based on a large number of soundsequences in a creative-generative way with consistent semantics leading (first) to vocabulary and (then) to grammar.

Soundsequences Used by Animals were Taught them by Humans. Parrots and some ape individuals learn human soundsequences that humans teach them. However, these animals are not able to teach themselves, and learn soundsequences intraspecifically. They are unable to create new sound sequences on the basis of the learned patterns. See also the Universal "Human Soundsequential Language is Intraspecific" below.

Soundsequential Language is Only One of the Many Communication Systems Humans can Handle. It is, for humans, certainly the most important, with an enormous storage value (stated already by de Saussure). However, it is neither miraculous nor unique; and it is not even the best that humans are able to design. Humans can handle all subhuman communication systems. However, most of these subhuman systems are "excess baggage" for "homo soundsequentialis".

Soundsequential Language is an Invention of Cro-Magnon Man. This means that Human Language cannot be older than about 30,000-35,000 years. It also means, however, that man lived "99 % of his phylogenetic history" (about four million years) without (soundsequential) language (Décsy 1985.270). Neanderthal man may have vocalized, perhaps he was able to produce also i and u. But he had no soundsequence (duality of patterning, see below Hockett's list). As soundsequence makes speech (human language), we can also state that Neanderthal man was unable to speak (to use a kind of soundsequential human language).

A Subsystemic Manner of Patterning in Soundsequential Language is Universal. With reference to Louis Hjelmslev, Hockett differentiated between Cenematic (Phonemic) and Pleremic (Morphemic) subsystems of patterning ("Duality of Patterning". UL 12). Today, Sydney M. Lamb (founder of stratificational grammar) operates with four strata: Semology, Lexology, Morphology, and Phonology ("Quadripartitiveness of Patterning" – see Makkai S 4.418, GLC 107). Such patternings are rather arbitrary settings of grammarians; nevertheless, they can in certain cases have universal (cross-linguistic) validity.

Evolutionary/Hierarchic Buildup of Grammar. The sound (human communicative noise produced in the lung-mouth channel: originally a kind of interjection) is older than the syllable. The syllable is older than the (underived, simple) word. The underived word is older than the compound. The compound is older than the derived word. The noun is older than the verb. The syntagma (word construction, like white house, a kind of compound or phrase) is older than the sentence. The simple sentence is older than the period.

The Lower the Evolutionary/Hierarchic Status of a Particular Part of Grammar, the Higher the Probability that it Contains Universal Similarities. In phonetics (phoneme system, sound production, syllable structure; an archaic part of the grammar), there are only minimal differences between languages. The number of differences grows in increasing order in –> the soundsequence structure (vocabulary), –> derivation, –> morphology, –> syntax (unless late intellectual contacts cause synchronizations in the latter categories, as in most modern languages in Europe due to the Latin-Greek-English syntactic traditions).

All Languages of the World are Alike. This was recognized by the Franciscan monk Roger Bacon (?1214-1294) in England, who stated that "Grammatica una et eadem est secundum substantiam in omnibus linguis, licet accidentaliter varietur". In Charles Ferguson's translation: Grammar is one and the same in all languages [of the world] in substance, though it may vary in accidents (S 1.9). In modern terms: in the deep structure (semantics), all languages of the world are alike; only in the surface structure (phonetic, morphemic, syntactic representation, lexemic representation) are there differences between them.

Translatability is Proof of the Identical Deep Structure of the Different Languages of the World. Were the languages of the world not alike in their deep structure, we would not be able to translate from one language into another. Translatability is the most striking evidence of the internal structural identity of the languages of the world.

Perfection in Competence is no Requisite for Communication. Stated by Saul Levin in a letter to Decsy (August 19, 1985). No speaker of a given linguistic community is actively in command of the entire Competence Code of the language he speaks. Competence commands are relative and gradual; at the top is the (basically, always relative) perfection in the mother tongue, on the bottom the pidgin-like command (broken or fragmentary knowledge) of a non-vernacular.

Human Soundsequential Language is Intraspecific. Sebeok K 571. If nonhumans perceive (parts of) the soundsequential human language, it is the result of human mediation (Garvin K 5). Meaningful soundsequences have always been taught by humans to animals (parrots, apes); animals cannot learn soundsequences with clear-cut consistent semantics from each other (i.e. intraspecifically). Exceptions are some acoustic signals (calls) of birds and other animals; however, these signals lack syntax, and are closer to music (displaying, in some cases, especially in birds, melodic/tonic elements) than to organized human speech (see Hartshorne 1973.18). See also the Universal "Soundsequences Used by Animals were Taught them by Humans" above.

Hockett's Feature List of the Soundsequential Human Language. Elaborated by Hockett (UL 1-13; cf. also HC 1982.6-8). These are: **1**. Vocal-Auditory Channel. **2**. Broadcast Transmission and Directional Reception. **3**. Rapid Fading (also called "Transitoriness, Evanescency"). **4**. Interchangeability (all normal communicators are interchangeably transmitters and receivers). **5**. Complete Feedback (we hear what we say: ability to repeat, store, and/or reproduce thoughts). **6**. Semanticity (denotationality). **7**. Specialization (we use speech only if it makes sense). **8**. Arbitrariness (iconicity is not a constitutive element of sign production; it never has been). **9**. Discreteness (pin vs. bin). **10**. Displacement (remembering things remote in space and time = memory). **11**. Productivity/Openness (ability to recognize new facts, generate new statements, and create new messages). **12**. Traditional Transmission (the conventions of language are acquired, basically, by teaching and learning; not inherited, not innate). **13**. Duality of Patterning (the same sound can have different value in different positions in the sequence, e.g. team vs. meat, Hungarian bök 'to strike, to kick' vs. köb 'cubic'). **14**. Prevarication (our messages are from a scientific point of view neither exact nor exhaustive). **15**. Reflexiveness (we can speak about language, but the bees cannot dance about dancing: they dance only about sites). – There are some differences in the number of features and their names between the two versions of Hockett's article (UL 1-13, HC 1982.6-8).

Every Language Changes in Time and Space. There are no petrified languages. Speech (as the surface projection of language) changes faster than the underlying language structure, the guiding principles of which have remained stable from the beginning of the emergence of soundsequential communication.

There are Three Basic Types of Linguistic Change. **1**. Addition (generates and introduces new features). **2**. Deletion (removes features). **3**. Replacement (exchange of the place of existing features). Reference is made to H. M. Hoenigswald in LU 50.

Every Soundsequential Language has Empty Slots (Gaps). This reflects structuralists' views that there is a large matrix with a sign (possibility) at every point, and that practically only a (small) part of (theoretically) possible complete matrix is exploited for positive sign setups in communication (Gamkrelidze S 2.35). Identifiable empty slots can play a role in opposition buildings, e.g. English diskø vs. disks (ø is here the sign for an empty slot or zero morpheme). For the concept zero morpheme see the excellent book by Meier 1961.

Every Language has Zero-Morphemes if we set the number of signs higher than the actual sign inventory practically offers. The Zero-Morpheme can have many different variations according to the paradigmatic set in which it occurs. For example, it is the sign of singular in the opposition houseø vs. houses, but it is the sign of the non-3rd person singular indicative present tense in (they) talkø vs. (he) talks.

Soundsequence Languages Actually Use Only About 30-40 Percent of the Theoretically Possible Sign Inventory. 60-70 % of the possible sign inventory remains unknown/unused in every language. Languages which do not use, for instance, affricates, tone/stress/quantity correlation, glottalized or labialized consonants, etc. make no use of existing possibilities of sign creation.

Binarity/Dichotomy is a Basic Structural Feature of Soundsequential Language. By interpretational techniques, triadic systems can be broken down into binarity subsets. For example, the triadic set /a/ vs. /o/

vs. /u/ can be viewed as /a, o/ vs. /u/. The vowels /a/ and /o/ constitute a phonogenetically and structurally affiliative unit as opposed to /u/. In the triadic Personal Pronoun system /I/ vs. /you/ vs. /he, she, it/, the 1st and the 2nd person constitute as a subset one unit as opposed to the 3rd person.

2.1 Phonetics/Phonology

Phonetic and Semantic Terms are by Definition Universal and not Specific (Saporta UL 63). This refers, apparently, to the earliest ages of the development of human language.

Clicks in Soundsequential Use are Restricted to Khoisan (according to Ruhlen 1977.140. See also GLC 66-67, as well as the basic publication of R. Stopa: Schnalze – Ihre Natur, Entwicklung und Herkunft. Bloomington: Eurolingua 1986). It is unclear whether the efferences of the brain transduced through nervepaths to the articulatory organs (larynx, velum, tongue, lips) cause direct noises ("clicks") at their arrival at the executive points (reference points) or just prepare the position of the particular organs for the air stream coming from the lungs, which then makes the noise (sound). For review of the problem see Lass1982.1. The Khoisan clicks may be affect products which received soundsequential acceptance only late in this phylum (which was spread before 2000 B.C. throughout the entire southern third of Africa).

Clicks cannot be the Forerunner of Human Sounds. The human sound is per definitionem an expiratory product made by the air which we press from the lungs through the larynx-pharynx tube and the oral cavity (occasionally in interaction with the nasal cavity). Clicks are noises created without pulmonary air pressure either by mechanical motion of the tongue (in interaction with lips and teeth) or by the lips alone. They also can be produced by (non-pulmonary) air previously accumulated and then released in the mouth cavity (ejectively). Also, the implosives are not real sounds even if they resemble the real p or t etc. to a certain degree, they do not meet the basic requirement that a human sound has to be produced by pulmonary air pressure leaving through the mouth cavity. Therefore, clicks are not sounds per definitionem. They also cannot participate in quantity, tone, or stress correlations, which are supposedly the most archaic vehicles of sound production. Clicks are not variable like pulmonary sounds.

Sounds/Phonemes

Sound is First of All Gesture/Motion. Recognized by Plato and repeatedly stressed by W. Wundt (for details see Décsy 1983.102). The (easily recognizable) "motion character" of sound seems to be a human feature. The motion of the lips and tongue as an articulatory procedure is only-human (the other primates are unable to move their lips and tongue in noise-making in the lung-mouth channel). The sound itself is a result of the air flow through the lung-mouth channel and the motion of the larynx, tongue, mouth, lips, etc.

Sound is, as an Articulate Unit of Speech, Human-Species-Specific (Only-Human). "The use of sound in the wider scheme of biological existence (i.e., outside the human realm) is rather uncommon. The overwhelming majority of animals are both deaf and dumb." (Sebeok K 573-574). A consistent and sophisticated noise production by the lung-mouth tube (Ansatzrohr) is, basically, avial and mammal. Avial noise production (bird singing) by lung-mouth is more developed than the noise production of non-human mammals. However, no species has ever reached the same "timbred differentiation" in (noise) production by lungs-mouth tube as humans. Only the timbrically differentiated human lungs-mouth noise product can be called "sound". A sound is, according to this, an articulated (timbric) lung-mouth noise with clear-cut and socially valid semantic value common only to human societies.

Only (Important Sounds) Phonemes are Perceived. "It is well-known that we are able to perceive minute phonetic differences when they have a linguistic function, and we neglect great differences between sounds if the language does not utilize them" (Balla 1984.42). This is one of the important facts which make the existence of phonology besides phonetics compelling.

Four Basic Features of Human Vocalic Sound. These are the stress (dynamic accent, strength), the tone (melody, pitch, high-low scale), the length (quantity, duration), and the timbre (color, Klangfarbe, Tonespecifics). Stress is certainly the most archaic basic feature, followed by tone, length, and timbre (in this evolutionary setting). The first three (stress, tone, and length) have their nerve centers in the cerebellum), and are certainly subhuman or even (in their predispositions) submammal. Timbre is only-human, having its neural center in the neopallium (cortex). Broca's and Wernicke's areas in the dominant (regularly, left) hemisphere of the brain are responsible only for the timbre (and timbred vowel combinations as well as their semantics), and not for the stress, tone, and length of the sounds or words. Supposedly, the untimbred vowel-consonant ("vocsonant") ə/ʔ (and its separated secondary realizations, the vocalic ə as well as the consonantic stop ʔ) are also centered in the subcortical areas of the brain (like stress, tone, and length).

Timbrically, Bird Song Is Not Variable; Nevertheless, It May Have Stimulated Human Sound Production. In terms of human language, the song of birds is timbrically monotonous: it is, mostly, i-like ("whistling"), sometimes u-like. It may be variable in pitch, but not in timbre. (cf. Hartshorne 1973.49). Nevertheless, bird song may have been a stimulus for man, in the early stages of human language development, to learn how to make an i-like vowel.

Stress Can be Used for Semantic Contrasts Only in non-Monosyllabic Languages. In monosyllabic languages it can not have distinctive functions. See English contact (noun) vs. (to) contact (verb); Komi-Permyak olan 'you live' vs. olan 'the life'.

Positionally, There Are Three Types of Word Stress: stable, mobile (can be applied only in non-monosyllabic languages), and facultative. Komi-Ziryan and Erza-Mordvin have, e.g. facultative stress. Details D 202. Only mobile stress is used for semantic differentiation.

Intensity Differences between Stressed and Unstressed Syllables are Taxonomically Relevant. There is a "cumulative" (peak) stress appearing in "stressed" syllables which is very strong compared with the stress of unstressed syllables (e.g. in English or Russian). The other type can be called "distributive stress" (non-peak stress, Verteilungsakzent), which is relatively less intensive compared with the stress of the unstressed syllables. This type of stress is common in German and French. also in Finnish and Hungarian. These stress types represent a characteristic feature of the articulatory basis; and are of special importance in foreign language teaching. D 202.

Languages with Cumulative Stress Have Reduced Vowels. Good examples are English and Russian.

Languages with Distributive Stress Have no Reduced Vowels. Most Western Slavic languages, German, Italian, and most Finno-Ugric languages (except Yurak and Mordvin) belong to this stress type.

Every Language Makes Use of Distinction of Vowel Color (Timbre). Vowel color (timbre) is defined in the experimental phonetics as combination of formants (Hockett UL 25).

Molnár's Four Noise Product Types. Fixed pitch and no timbre is music, free pitch and fixed timbre is human speech; fixed pitch and fixed timbre is human song; unbound pitch and no timbre are all other noise products of the world. Stress and length are irrelevant in this classification. (Details Décsy 1977.77).

Accessorial Features of Phonemization. These add in each case archaic elements to the customary articulation of a sound in order to produce new full phonemes. The accessory features are palatalization (based on the addition of a slight j), velarization (adds a kind of k/q element), pharyngealazation, laryngealization, labialization (utilizes a u/w element), aspiration (h added), nazalization, lengthening (leading to quantity correlation), "stressing" (using change in placement of stress in inflectional or derivational sets), and tonization (adds tones). Phonemizations with these accessory features seem to be late and secondary (even if the stress, tone, length as differentiating features in themselves are primeval and archaic); they are used because the customary timbre-vowels or the regular consonant sets (simple stops, occlusives/fricatives/spirants, nasals, lateral and rolled consonants) do not suffice for the production of distinctive sound sequences for new words. Tonemes, chronemes (opposition of short/long, quantity correlation), and stressemes (cf. contact and contact) are sign elements from the pretimbric times which were dropped as "excess baggage" when timbres were introduced, and were reintroduced as useful primitive differentiation in our modern languages.

The Phonetic Efficiency of Languages is Distributed Roughly around the 50 Percent Point. "There are sets of pressures bearing on any phonetic system which cause it to maintain some optimal efficiency value". After Greenberg UL.XVII-XVIII.

50 Percent Redundancy is a Regular Case in the Languages of the World. This follows from the previous. Cf. Hockett's formulation: "In every human language, redundancy, measured in phonological terms, hovers near 55 %". (See UL 24.)

A Historical Tendency Toward Phonological Symmetry is Universal (Hockett UL 25 with reference to Jakobson). Hockett stresses that a perfect symmetry can never be reached. It is difficult to assume that such a teleological principle would be the primary cause of the expansion of phoneme systems (by addition). Hidden factors should be sought in cases in which this principle seems to apply. The purpose of speech is communication and not the establishment of balanced systems in a visual set of geometric formations.

Binarity is the Ruling Principle in the setup of oppositions (correlations) in phoneme systems. Triple correlations are rather uncommon. In any case, the vowel sets a/o/u and ä/e/i (existing in many languages) constitute triplets. The partial sets u/i/ü and e/o/ö or œ/ö/u are also triadic. In Estonian, there seems to be a triadic durational (quantity) correlation in both the vowels and the consonants (see below). The Classical Arabic (and Proto-Semitic) vowel system is binary: /a/ is opposed to /u/ and /i/. In terms of evolutionism, this system is more archaic than the supposed (triadic) Indo-European. In Morphology, the system of personal pronouns constitutes a triadic system (see below).

The Oldest Articulated (Timbric) Sound Products of Humans were Undivided Vocsonants. These consisted of i/j, a/h and u/w. (Décsy 1985.266).

Vowels

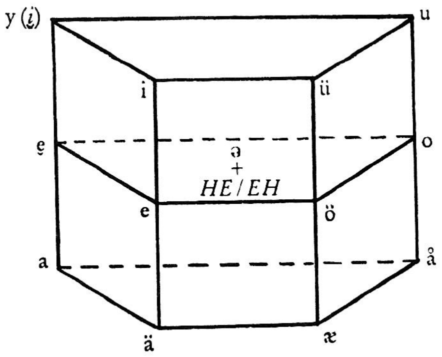
A Language Never Has a Greater Number of Phonemic Contrasts in the Vowels of Unstressed Syllables than it Has in Stressed Syllables. (Ferguson UL 54, after Trubetzkoy). However, this may be the result of secondary vowel simplifications (coincidences) in the unstressed syllables (see reducing languages below).

Long Vowels (also Diphthongs) and Tone Occur as a Rule only in Stressed Syllables. This is clearly the case in Finnish in connection with the diphthongs. For similar phenomena in Spanish cf. Sol Saporta (UL 63 and 67). Follows from the previous.

There Is no Proven Univocalism or Bivocalism in the Present-day Languages of the World. All languages of the world have at least three vowels: /u/, /i/, /a/ (example: Classical Arabic, details GLC 121, Décsy 1983). Most languages have, as a rule, five vowel phonemes in a triangular pattern: /u/, /i/, /e/, /o/, /a/.

The Maximal Number of Vocal Phonemes is Twelve. The 12 vowel phonemes can be arranged at the corners along the edges of a cube (according to Forchhammer). As an example, I give here the 12-member phoneme system of vowels in Selkup (according to Décsy 1977.56). The timbreless vowel /ə/ appears in the center of the cube. For more see John

Crothers (S 2.93-102. chart p. 134: most of his 12 descriptive statements based on a corpus of 209 languages cannot be regarded as universals).



The Most Common Pattern in the Languages of the World is the Three Level, Five Unit Vowel System of Classical Latin. The system comprises /i, u, o, e, a/, or as a diagram in triangle shape:

i u

e o

a

Slightly more than 25 % of the units in Ruhlen's corpus (about 700 languages) exhibit this pattern. Breakdown according to continent zones: Oceania 40 %, North America 33 %, Europe 23 %, Asia 22 %, Africa 19 %, South America 15 %. This was the system of Proto-Indo-European, and has been maintained by many separate languages of this phylum. The five-vowel system extended by schwa (occuring mostly as variation in unstressed syllables) is to be found in 7 % of the corpus (Ruhlen 1977.147-148). The system represents a three-level pattern and thus does not meet the requirements of the dichotomy. Therefore, it must be secondary. It seems that /o/ and /e/ in Indo-European were originally only positional variations of other vowels.

A Two-Level Triangle is Found in 7 Percent of the Languages of the World. The system comprises /i, u, a/, see Ruhlen 1977.148, in diagram:

i u

a

Its main carriers are Classical Arabic and the archaic forms of the Indo-Iranian languages. This may have been the Pre-Proto-Indo-European system too. The pattern is dichotomic, therefore, it must be very old. It is also simpler than the five-unit system with /i, u, o, e, a/; this is another feature confirming the archaism of the system/i, u, a/.

Square Vowel Systems Exist only in Languages which Have/Had Vowel Harmony or Umlaut. In diagram the system looks like this:

i u
e o
a a

This is the typical vowel system of the Uralic languages. Note that a is the conventional sign for open e (ɛ); same as a in English bad.

Front Rounded Vowels Occur in 12 Percent of the Languages of the World (stated after Ruhlen 1977.148). Asia 35 %, Europe 30 %, Africa 3 %, Oceania 3 %, North America 1 %, South America 0 %. Note that the concentration areas of the front rounded vowels are Europe and Asia (Eurasia). They are typical elements of the Uralic, Turkic, and Mongolic languages.

Front Rounded Vowels are Late and Secondary in all Languages in which they Exist. This can be proven with the method of historical grammar (e.g. for the Germanic languages and for Hungarian) or with the method of comparative grammar (for the Uralic languages). The insistence of some Altaists that Proto-Mongolian or Proto-Turkic phoneme systems included u or o is erroneous.

10 Percent of the Languages of the World Exhibit Back Unrounded Vowels. Most of them are in Asia and South America (a rather unusual distribution), cf. Ruhlen 1977.149: Asia 30 %, South America 21 %, North America 7 %, Europe 5 %, Oceania 4 %, Africa 2 %. "While most languages with back unrounded vowels also possess back rounded vowels, in South America it is not unusual for a language to have" unrounded u /ɯ/ or unrounded o /ɤ/ without the corresponding rounded vowel. Ruhlen 1977.149.

22 Percent of the Languages of the World Utilize Nasalized Vowels. Most of them are in South America (52 %); Africa has 29 %, North America 24 %, Europe 21 % , Asia 19 %, Oceania 3 %. (Ruhlen 1977.149).

Voiceless Vowels are Rare. However, they exist in Lapp and in Finnish. Practically, they are interpreted as a connection of a vowel + h, e.g. Finnish lahti 'bay, gulf'.

Long Vowels

48 Percent of the Languages of the World have Long Vowels. According to Ruhlen 1977.150: 62 % in North America, 53 % in Asia, 50 % in Africa, 46 % in Europe, 43 % in Oceania and 26 % in South America. As lengthening is a secondary (accessorial) feature of phonemization, the South American low-level data may represent a retention (archaic feature).

Vowel Length Often Interacts with Tone and Stress. Rising tone may lengthen the vowels. Maddieson after Gandour S 2.355. In Russian and French, vowels of stressed syllables are generally longer than in unstressed syllables.

Triadic Vowel Length Contrasts are Rare in the Languages of the World. Estonian (except for its northeastern dialects) has this distinction. Triadic vowel quantity correlation has been supposed also for Lapp. Here some Estonian examples will be offered as illustration:

> jama nom. sing. 'nonsense' (first a short, second a half long)
> jaama gen. sing. jaam 'railroad station' (first a long, second a half long)
> jaaama illative sing. of jaam 'railroad station' (first a over long: aaa, the a of the second syllable is short)
>
> sagi 2nd pers. sing. imper. sagima 'to pull' (a short, i half long)
> saagi gen. sing. of saak 'booty, plunder' (aa long, i half long)
> saaagi illative sing. of saak 'booty, plunder' (aaa over long, i short)

There is some formal and combinatory limitation of occurence. 1. Triadic sets occur only in the first syllable of non-monosyllabic words (non-monosyllabic in inflected forms; they can be monosyllabic in the basic form – see saak in nominative above); they never occur in monosyllabic form sets. 2. The over long form occurs only if the second syllable is short. Historically, the second syllable is short when its antecedent syllable was open (jaaama < jamahan < jamasan). 3. Short and long in the first syllable occur only if the final vowel is half long; the final vowel is half long when the antecedent syllable was either open (then the short grade occurs: jama 'nonsense' in nominative) or wen it was closed in an inflected form (jaama: gen. sing., originally jaman; saagi gen.sing. of saak 'booty, plunder', originally saakin). Phonological evaluation of the phenomenon in L. Posti, Journal de la Société Finno-Ougrienne 54:2(1948-1950).8. The triadic sets are limited, further, to certain grammatical categories (genitive/illative, 2. person imperative).

The Quantity Correlation of Short to Overshort is Rare. Nevertheless, it exists at least in Ostyak (see Décsy UAJb 32.164-174, 1960), in Karok (a Hokan language, R 218) , in Kurdish (R 228),and apparently, also in Hindi (R 192) and East Aramaic in Northwest Iran (R 160). For symmetry's sake, the overshorts are often interpreted as short, and the shorts as long. The overshorts are also called "reduced" (apparently, in quantity). However, the term reduction should be reserved for quality (timbre) losses only (see UAJb 32.164-172).

Diphthongs

Diphthongs are Tautosyllabic Connections of two (Mostly Different) Vowels. For Indo-European, the following diphthongs can be assumed (Szemerényi 1970.35): /ei, oi, ai/ and /eu, ou, au/.

Diphthongs Exist only in Languages which Have or Had Long Vowels. Details Décsy 1970.15.

In Diphthongs, the Archaic (Semi-) Vowels i and u Commonly Appear. See the Indo-European examples above. However, i and u as diphthong components can be interpreted as (consonantic) /j/ or /w/. Such monophthongic/consonantic interpretations are common.

Non-u/i-Diphthongs are Rare and New. See French Pierre (< Per < Peter), mois (< meum, or, mensis); Finnish mies 'man' (< mees), yo 'night' (< öö). This Finnish sound change (diphthongization) took place in the 16th century (or shortly before). There are phoneticians who regard only the non -/i, u/ diphthongs as real diphthongs.

Diphthongs and their Number are Important Taxonomic Indexes in the Languages of the World. There are monophthong- and diphtong-oriented languages. English, Russian, and Hungarian are monophthongic languages (at least in their acrolects); French, Italian, the Balto-Finnic languages, and Lapp are diphthongic. The monophthongic type prevails; an estimated 70 % of the languages of the World are monophthongic.

Diphthongs can Develop from Vowel + Consonant Connections. For example Finnish sauna < *sakna 'steam bath'), Finnish kaula 'neck' < *kaklas (cf. Latvian kãklas 'id.'), seura 'society' < *sebra 'friend' (cf. Latvian sẽbras 'id.').

As a Result of Monophthongization, Diphthongs Often Develop into Long Vowels. Cf. the English pronunciation of such words as August, Sauna, Pierre (place name), bowl.

Consonants

The Number of Consonant Phonemes in a Language Varies Between Seven and about Sixty. However, the number of consonant phonemes in most languages does not exceed 20 to 30. Hawaiian (Austro-Tai/Austronesian, about 300 speakers) has only seven consonants (p, k, ?, h, m, n, l; – Ruhlen 1976.204). For Lak (a Caucasian language in Daghestan) about 45 have been listed, for Lapp (a Uralic language) about 90 consonants (Ruhlen 1976.229,and 230), obviously because many positional and accessorial variations (long consonants, palatalized consonants) were counted as independent phonemes. The Finnish phonetician Franz Äima (1875-1936) operated with about 800 Lapp sounds (registering vowels, diphthongs and consonants with all their variations).

The Consonant /h/

The Consonant h Constitutes a Class of its Own. This conclusion has been reached by the Finnish phonetician Kari Suomi; he set up (beside the obstruents /ptks/ and the semivowels /vj/) a separate "glottal class" for the Finnish /h/. This pure phonological/phonotactic analysis strengthens the views on the high chronological priority of h over the resonants (see Suomi 1985).

Roughly 35 Percent of the World Population does not Apply h in their Mother Tongue (data by Décsy in GLC 17). It is surprising that, according to Ruhlen, 36 % of the languages of the world do not include h in the phoneme inventory (GLC 17). From here it follows that about

Two Third of the World Population Utilizes the h in their Phoneme System. The h is, however, in many instances not an old preprotolinguistic element (hP) in the present-day languages but a secondarily introduced phoneme (hN) which developed from s, k, t, kh, g, etc.

Glottal Stop

Glottal Stop Occurs in roughly Half of the Languages of the World. Stated after Ruhlen 1977.139 (corpus of about 700 languages). I suppose that a "hidden glottal stop" exists in many further languages; however, it is not realized by traditional script and remained unnoticed by phoneticians (see my remarks in GLC 29).

It is Hard to Determine the Sonority Grade of the Glottal Stop. As a unique feature, Yurak (Nenets), a Uralic language, has both voiceless and voiced glottal stops. The voiced glottal stop of Yurak (written in the official orthography as /"/) came into being from voiced consonants /n, ŋ, j/, with which it alternates. The voiceless glottal stop /'/ alternates with the voiceless consonants /t, s, k/. Details Yurak 72.

Stops

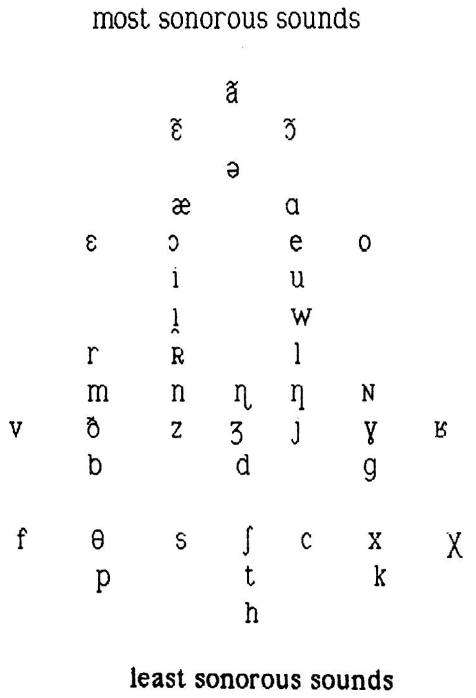
Stops are Older than Fricatives both Phylogenetically and Ontogenetically. Stated by R. Jakobson 1949, – see Gamkrelidze in S 2.32. There are some serious restrictions to this, as w, j and probably the nasalized n (ŋ) are older than the stops /p, t, k/. See below.

10 Percent of the Languages of the World have Prenasalized Stops. In Africa 26 %, and in Oceania 22 % of the languages utilize prenasalized stops; nine Amerindian and two Asian languages also have them. Stated after Ruhlen 1977.140.

72 Percent of the Languages of the World Exhibit Sonority Correlation at Stops. Stated after Ruhlen 1977.140-141. The feature appears in Africa with 97 %, in Europe with 92 %, in Asia with 84 %, in Oceania with 63 %, in South America with 53 % and in North America with 46 % (distribution data according to Ruhlen 1977.140-141). Lack of the sonority correlation is regarded generally (and, particularly at the stops) as an archaic feature. It is surprising that this archaic feature (i.e., the lack of the sonority correlation) is maintained mainly in the indigenous languages of the Americas as in North Siberia.

The Least Sonorous Consonants are the Laryngeals, the Most Sonorous Consonants are w and j. Stated after Paul Dieth, see GLC 102.

It is no accident that the two most archaic consonant groups (laryngeals vs. w/j) are at the extreme opposite ends of the sonority set. The vocalic variations of w/j, namely u and i, are the least sonorous vowels (as opposites to a). See the diagram (sonority increases from the bottom to the top):



13 Percent of the Languages of the World Utilize Aspirated Stops. Stated after Ruhlen 1977.141. However, as Chinese, English, and German also use aspirated stops, about 30 % of the world population pronounces the stops p, t, and k aspirated.

Aspirated Stops Occur Only in Languages which have (or had) an h. After Hockett UL 25 with reference to Jakobson.

Fricatives

Almost All Languages of the World have Fricatives – at least one, but as a rule more than one. From the approximately 700-language corpus of Ruhlen, only 36 did not have fricatives (22 in Australia, seven in New

Guinea, three in South America, two on Bougainville, one in Africa, and one on the Gilbert Islands. Ruhlen 1977.141.

Pharyngeal Fricatives Occur only in 7 Percent of the Languages of the World. Stated after Ruhlen 1977.143. They are typical in the Afroasiatic and Caucasian languages. Pharyngeal fricatives are often regarded and counted as laryngeal (glottal) h (especially if there is no h in the particular language). Note that pharyngeal fricatives exist (according to rather common interpretations) in Ukrainian and South Russian. They may have come to this area from the Caucasus (or, perhaps, partly also through Yiddish mediation).

Affricates

The Affricates are Late and Secondary in Every Language. Experience shows that the four commonly used affricates (c/dz, ch/dzh, pf/bw) emerged late in languages which utilize them as phonemes or variations. Their early occurence in infant babbling is occasional (maybe a misinterpretation of s or kh) – see Reich 1986.

The Affricates are Consonant Diphthongs. They are connections of a stop and of a homorgan spirant: c < t + s, dz < d + z, ch < t + sh, dzh < d + zh. Homorgan means: articulated approximately at the same place.

Affricates Occur in 70 Percent of the Languages of the World. Stated after Ruhlen 1977.141. North America 90 %, Europe 86 %, Asia 82 %, South America 78 %, Africa 60 %, Oceania 28 %. As the affricates are evolutionarily young, their high percentage is somewhat surprising. Ruhlen's data are based on a corpus of about 700 sample languages; research of the 2,800 languages would lead, probably, to a different result. Note that the archaic indigenous language areas of Australia and South America display limited representation of the affricates. In languages with affricates, the affricates are often subject to limitations in occurence.

There are Evolutionary Status Differences between the Different Classes of Affricates. The consonants c (ts) and ch are probably more common than dz and dzh. The Scandinavian (North Germanic) languages have no c and dz (this is also in English where these affricates only occur in cases like cats and kids). It is remarkable that Acrolect French has no affricates at all (earlier ch > sh in pronunciation). This may be one of the "Scandinavian" (Norman, Viking) features of this language.

Labial Affricates are Rare in the Languages of the World. See the Standard German pf in Pferd 'horse' (very often pronounced as fert).

Liquids (Laterals and Vibrants)

Laterals Occur Commonly in Languages around the World. Stated by Ruhlen 1977.143. Distribution of these l-type consonants according to continent zones: Europe 100 %, Asia 96 %, Africa 89 %, North America 85 %, Oceania 73 %, South America 30 %. Note the lack in many of the languages of the archaic indigenous territories Oceania and South America.

77 Percent of the Languages of the Word Utilize Vibrants. Vibrants are the r-type consonants. Breakdown: Europe 100 %, Africa 89 %, Asia 83 %, South America 81 %, Oceania 75 %, North America 45 %.

5 Percent of the Languages of the World is Liquidless (i.e., have neither l nor r). Stated after Ruhlen 1977.144. "All but one (Efik) of the liquidless languages are found in Oceania (the Indo-Pacific phylum), North America (chiefly the Central Algonquian and Siouan languages), or South America. (Ruhlen 1977.144).

The One-Liquid Languages are Bearers of Significant Evidence with Regard to Language Origins. They carry on a monolinear realization of the supposed (pre)protolinguistic undivided liquid phoneme L/R. Chinese has (in a simplified formulation) only laterals (l), and Japanese only vibrants (r). The numerous alternations between l and r in Indo-European and Uralic etymologies may point to an undivided liquid consonant L/R in the protoforms of the two phyla. GLC 93.

Retroflexes

Retroflex Consonants Occur Typically on the Indian Subcontinent and in Australia. See Ruhlen 1977.145. They are found, sporadically, in other continents too. Their occurence is often restricted to certain sets. 19 % of the languages in Ruhlen's corpus exhibit retroflex consonants.

Implosives

Implosives Occur in Only 10 Percent of the Languages of the World. They represent an areal characteristic of Sub-Saharan Africa, Southeast Asia, and Southern Mexico. Africa 35 %, North (Central) America

10 %, South America 5 %, Asia 5 %, Oceania 5 %, Europe 1 %. After Ruhlen 1977.147.

Consonant Accessorials

There are at Least Eight Types of Consonant Accessoriality: Length, Palatalization, Labialization, Velarization, Glottalization, Pharyngealization, Aspiration, and Laryngealization. Except for the length, they all add a slight element of a different acoustic impression than the main consonant exhibits.

There are Three Kinds of Palatalization: Tongue-fronting, tongue-raising, and spirantization. They can occur either individually or in different combinations (D.N.S. Barth in S 2.47).

Long Consonants

14 Percent of the Languages of the World have Long Consonants. Stated after Ruhlen 1977.145. Europa 28 %, Africa 27 %, Asia 14 %, North America 7 %, Oceania 5 %, South America 1 %. Quantity correlation of consonants is not a wide spread feature. It may be secondary. For Proto-Uralic or Proto-Indo-European, no long consonants can be supposed. Out of the approximately 25 Uralic languages, today only Hungarian, Finnish, Estonian, and Lapp exhibit impressive inventories of long consonants. Except for Hungarian, the quantity correlation is restricted in the Uralic area to certain groups of consonants (mainly to stops and liquids) and to certain positions (internal syllable boundaries) or alternations. The near absence of long consonants in South America seems to be a retention.

Total Quantity Correlation of Consonants is Rare. The only European language in which all consonants have a long variation is Hungarian. The long consonants in this language can occur only in intervocalic and in final position, however – not in initial position or after another consonant (e.g. karddal 'with a sword' is pronounced as kardal). Nevertheless, Hungarian seems to be the most long-consonant-intensive language of the world. All long consonants represent an innovation in this language, they began to spread there after the 14th century. Telugu also seems to have a total quantity correlation of the consonants (see R 276).

Triadic Quantity Correlation of Consonants is Rare. An exception is Estonian (all dialects except for the Northeast, which is close to Votic and

Finnish). However, three grades never occur within the same paradigm; i.e., the three sequences are cognates. Illustration:

lina (i short, a half long, n short) 'flax, cloth, linen'
linna (i short, a half long, n long) gen. sing. of linn 'city'
linnna (i short, a half long, n overlong) illative. sing. of linn 'city'

Additionally, there is some formal and combinatory limitation. 1. Triadic sets occur only in the first syllable of non-monosyllabic words (non-monosyllabic in inflected forms; they can be monosyllabic in the basic form, see linn in nominative above); they never occur in monosyllabic forms. 2. The overlong form occurs only if the second syllable is short. Historically, the second syllable is short when its antecedent syllable was open (linnna < linnahan < linnasan). 3. Short and long in the first syllable occur only if the final vowel is half long; the final vowel is half long when the antecedent syllable was either open (then the short grade occurs: lina 'flax, cloth' in nominative) or closed in an inflected form (linna: gen. sing., originally linnan). The triadic sets are further limited to certain grammatical categories (genitive/illative, 2nd person imperative). – Phonological evaluation of the phenomenon in L. Posti, Journal de la Société Finno-Ougrienne 54:2(1948-1950).8.

Nasals

In Human Vocalization, Phylogenetically, Nasals are Older than the Oral Sounds. Oral articulation became possible only when humans acquired the ability to make the velopharyngeal closure (details in Décsy 1985.267). By velopharyngeal closure we separate the mouth cavity from the nasal cavity and this makes the air flow exclusively through the mouth cavity. In the "pre-velopharyngeal times" all sound products were nasalized, because the air left the larynx-pharynx area mainly through the nasal cavity, producing in this way nose color nuances. In some instances, the proto-sounds may have been neutral: neither nasal or oral or both at the same time. However, as the natural channel of the air flow from the lungs is through the nose (and not through the mouth), complete nasalation is more probable in the pre-velopharyngeal time than a naso-oral noise production.

All Languages of the World have at least One Nasal Consonant. It is the nasal n "produced by a complete oral stoppage (e. g. apical, labial), velic opening and by vibration of the vocal cords". Ferguson UL 56 termed the n Primary Nasal Consonant (PNC). A more specific formulation is:

All Languages have Nasal Consonants. Stated by Greenberg UL.XX. Exceptions seem to be three Salishian languages (Puget Sound 15, Quinault 10, and Twana 10 speakers) as well as Mura-Piraha (110 speakers, Brazil/Amazonas). This unusual exception requires clarification. See also Ruhlen 1977.143.

Languages with /n/ as the only Primary Nasal Consonant are Chiefly in the Western Hemisphere. E. g. Tinglit, some Iroquois languages, Arapaho. Ferguson UL 57.

There are no Languages with /m/ as the only Primary Nasal Consonant. Or: if a language has an /m/, it also has an /n/, but not vice versa. Follows from the previous.

In a Given Language, the Number of Primary Nasal Consonants is Never Greater than the Number of Series of Obstruents. Ferguson's formulation in UL 57.

Secondary Nasal Consonants occur only in Languages which have One or More Primary Nasal Consonants. Ferguson UL 58. There are at least six subtypes of Secondary Nasal Consonants: Voiceless Nasals (Mordvin, Danish, Lapp, Welsh, etc.), Aspirated Nasals (Marathi), Glottalized Nasals (Chontal/Oaxaca), Palatalized Nasals (Russian, Mordvin, Karelian), Emphatic Nasals (Syrian Arabic), Prenasalized Stops (Fiji), Nasalized Clicks (Zulu). Ferguson states (in UL 58) that these Secondary Nasal Consonants developed from clusters.

Nasal Vowels exists only in Languages which have Primary Nasal Consonants. Sample languages: French, Bengali, Taos (Ferguson UL 58). Nasal vowels always develop from connection of Primary Nasal Consonant + Vowel. Ferguson UL 59.

Nasal Syllabics, apart from Borrowings and Analogical Formations, always Result from Loss of a Vowel. Ferguson UL 59. Cf. Japanese /ñ/, Ewe and Khosa /m/.

Increasing Phoneme Numbers

Compensatory Alternatives to a Small Phonemic Inventory. There are four of them: **1.** The morphemes are longer than usual. **2.** Larger than average percentage of possible combinations. **3.** Larger number of homonyms. These three items were listed by Saporta (after Jakobson) in LU 63 and 72. We can add here as **4.** Accessorial (suprasegmental, prosodic) features of phonemization by stress, tone, length in the vowels; and by length, palatalization, glottalization, velarization, aspiration, pharyngealization, labialization in the consonants.

Sound Iconicity/Toconism

Sound Iconicity (Toconism) is Universal. "Regardless of the culture or language. The universal is not in any specific sounds associated with specific experience. It is that all languages seem to exploit sound symbolism [=sound iconicity] in some implicit way that is systematic and internally consistent." Jack Rosenthal, with reference to Steven Feld (Kaluli specialist; Kaluli is a New Guinea idiom called today Bosavi). New York Times Magazine, June 30, 1985 (Section 6), p. 8.

Smallness is Expressed in Word Stems: 1. by Front and/or High Vowels; 2. by High Tone or 3. by Various Kinds of Consonantal Ablaut. Cf. little, slim, thin, minor; Italian piccolo; Rumanian mic. Subgroups of smallness are diminutives, proximal local adverbs, and demonstratives. See Hungarian itt 'here'/ez 'this' vs. ott 'there'/az 'that'. Not valid for affixes. There are many adverse samples (English big, however, old bugia), small, Latin parvus. Although Greenberg (S 2.6) stresses the "lack of linguistic respectability" of sound symbolism (sound iconicity), the present universal (contemplated already by Plato in his Kratylos) can claim validity for the early stages of language origin. Ultan S 2.525, Jakobson 1979.178.

Bigness is Expressed by u. A favorite sound icon of Iván Fónagy (details Décsy 1981.18). Jakobson 1979.176 writes that a little German boy began to call his father "pupu" (instead of "papa") when the father wrapped up in a heavy fur coat. Compare also the sets dzharar 'to creep' (in general), dzhirir 'creep' (up to small beings) and dzhurur 'to creep' (up to big and frightful animals) in Batak (a West Indonesian language, old Batta, according to von der Gabelentz), see Jakobson 1979.179-180. The Batak data are insightful because the three archaic vowels /a, i, u/ denote here in a kind of Ablaut-set semantic differentiations. Striking examples are the so-called Vowel Harmonic Parallel Forms of the Uralic and Altaic languages – cf. Hungarian csúcs 'peak, summit' vs. csücs 'tip, corner (e.g. of a cushion'), Finnish tuhma 'naughty, bad' vs. tyhmä 'silly, stupid, foolish'. In both languages, the u-form expresses a higher intensity of quality/quantity than the front vowel variation.

Fastness Is Expressed by i. Examples German Blitz 'lightning' (Jakobson 1979.179, or English zip. Fastness may be a metaphor of smallness, therefore this is only a variation of the smallness-i-case (see above). This latter question remains, at least from a psychological point of view, open.

Words Designating Curvature Often Contain an r. Ascribed to J. Nyikos. Exemples: /turn/, /curve/, German /Kreis/ 'circle', Finnish /piiri/ 'circle', Hungarian /kör/ 'circle' (Décsy 1981.19). These types of sound iconicity may not be very old. As the r itself is a relatively young consonant, the words in question may be the results of a late autogenesis (German "Urschöpfung"). GLC 122.

The Consonant m Designates Round Objects. Stated by Georg von der Gabelentz (1840-1893) on the basis of observations concerning autogenetic (tone-iconic) trends in child speech. Jakobson 1979. 178. No claim of this principle for Language Origins times.

Laryngeals as well as m, ŋ are the Basic Phonic Material of Prearticulative Sound Products. Laughing, crying, coughing, and in some so-called "global words" (René Spitz's term) – cf.Latin mamma 'mother's breast', in Georgian mama 'father'. Helmut Gipper in K 544.

There is a High Probability that Words Designating the Female Parent will Have a Nasal. Greenberg LU.XII.

In the Word for Cuckoo a k Appears in Many Languages. English cuckoo, French coucou, Spanish cuclillo, Italian cuculo, Rumanian cucu, German Kuckuck, Greek kókkyks, Russian kukushka, Hungarian kakuk, Finnish käki, Zyrian kok. Ullmann UL 225. An onomatopoetic (sound imitating) word.

Verbs for Snoring Contain an r in Many Languages. English snore, German schnarchen, Dutch snorken, Latin stertere, French ronfler, Spanish roncar, Russian khrapeť, Hungarian horkol, Finnish kuorsata (see Ullmann UL 225).

Verbs for Whispering Contain an s, sh, or ch in Many Languages. English whisper, German wispern/flüstern, Norwegian hviske, Latin susurrare, French chuchoter, Spanish cuchichear, Russian sheptat., Hungarian susog/suttog. Ullmann UL 225.

The Consonant l Expresses Softness. Impressive examples from English, French, Finnish, and Hungarian poetry in Ullmann UL 226-227.

The Consonant r Expresses Aversion. The consonant r (and the l too) is often regarded as unpleasant. Lipograms (textes without r) are common in Russian and German. See GLC 93.

Evolution and Markedness

Less Marked Phonemes have Evolutionary Priority over Marked Ones. The least marked vowel is the reduced e /ə/, and the least marked consonant is the glottal stop /ʔ/. All other vowels and consonants can be derived by markedness additions (extensions) from /ə/ and /ʔ/, respectively. The phonemes /ə/ and /ʔ/ were undivided-neutral in their original appearance in the premarkedness times, i.e., they were at that time neither voiced nor voiceless ("voiceless media"), neither stops nor clusils, neiter nasals nor orals, but in a certain sense all these at the same time. For details see Décsy 1977.44.

Normal Hierarchy of Increasing Markedness in Vowels is /ə/ -> /a/-> /o/-> /u/ (back set); /ə/ ->/ä/ -> /e/-> /i/-> ü -> /ö/ -> /œ/ (front vowel set). For the specific phonogenetic placement of /a/, /i/, /u/, see the Proto-Vocsonant Universal below.

Normal Hierarchy of Increasing Markedness in Consonants is unaspirated -> aspirated-> glottalized. Gamkrelidze after Greenberg S 2.12. This is, of course, just one aspect of the markedness sets. See below.

Proto-Vocsonants are undivided proto-phonemes (neither vowels nor consonants, but in a certain sense both at the same time). They were sounds produced, without intensive participation by the lips and tongue, basically in the laryngeal area. It has been suggested to set up three such vocsonants: [a/ʔ], [u/w] and [i/j]. Their subsequent bifurcation led on one side to the vowels/a/, /u/ and /i/ (a classical archaic vocalic triad, cf. Semitic and Classical Arabic), and on the other side, to the archaic consonants /ʔ/ or /h/, /w/ and /j/, which are, in spite of their strong markedness, very archaic. The unmarked opposites /v/ to /w/ and /ç/ to /j/ are, therefore, always secondary (as experience proves). This is a very important exception to the developmental universal stating that the marked always has (chronological/structural) priority over the unmarked.

Markedness Status of h and of the Glottal Stop is Hard to Determine. Both "proto-consonants" are qualified, generally, as "unpairy voiceless". In some languages (Czech, Slovak, Ukrainian, Yiddish, Afrikaan) they are "unpairy voiced" (voiced because they originate in the voiced g). In some languages (Finnish, Hungarian) there is a voiceless and a voiced h (ɦ). In the Finnish loanwords of the Russian language (mostly in the North Russian dialects) the voiceless Finnish h appears as kh and the voiced as g (gh), e.g. lakhti 'bay, gulf' and raga 'money'. The markedness of the glottal stop in Yurak (Nenets) is controversial. Most native speaker grammarians of this language state that there is a distinctive voiced and a voiceless glottal stop in this language; they are so important for the native speakers that the official Yurak orthography introduced specific signs for them (/'/ =voiceless glottal stop, /"/=voiced glottal stop). Lability in determination of markedness is an additional evidence for the high evolutionary status of h and the glottal stop.

At the Stops and (Secondary) Spirants, Voiceless (Umarked) is Older than Voiced (Marked). As experience shows, this is always valid in the cases p/b, t/d, k/g, s/z, sh/zh, θ/ð etc. However, it is not valid for v/f, w/β, or for the nasal consonants, which are in their primary form ab ovo voiced. Indeed, in languages in which there is a voiceless/voiced opposition in the nasal or rolled consonants (as exist, e.g. in Celtic, Mordvin), the voiceless variation has been proven secondary. See below too.

Nasals, the Liquid Consonants, further /j/ and w (later › v) Are Ancient Orphan Primary Voiced Phonemes. Involved are /m/, /n/, /ŋ/, /l/, /r/, /j/ and /w/. However, their proto-forms may have been neutral concerning the sonority correlation ("voiceless media"): /M/, /N/, /ŋ/, /L/, /R/, /J/ and /W/. Capital letters indicate voiceless media, or (with some simplification) consonants which are neither voiced nor voiceless (neutral).

The Number of Marked Phonemes never Exceeds the Number of Unmarked Phonemes. Gamkrelidze after Ferguson S 2.12. Note that the same feature "can take the value of either a marked or an unmarked member depending on the concrete enviroment of distinctive features making up the content of the phoneme". Gamkrelidze S 2.21. This means that in extreme cases the analyzing linguist decides what is marked and unmarked; in the unavoidable arbitrations the frequency may play a major role. Voicelessness can be regarded, e.g., in case of necessity as negative markedness, and voicedness (e. g. in the primary nasals, liquids) as positive unmarkedness. Obviously, these are only terminological games made possible by change of the aspect places.

Stops are both Phylogenetically and Ontogenetically Older than the (Secondary) Fricatives. Roman Jakobson stated this in 1941 (see Gamkrelidze S 2.32). Children learn to pronounce secondary fricatives later than stops; in the language acquisition process they often substitute for fricatives with stops, e.g. Finnish puutta for puussa 'on the tree, on the branch of the tree'. (Secondary fricatives are: s, z, sh, zh).

Alveolar Fricatives are Newer than Dental Fricatives. Experience shows that s is older than sh, and that z is older than zh. The sh-lessness of Classical Greek or Latin is an archaic feature.

In All Languages which have Vowel-Harmony, the Vowel ä Exists. Details Décsy 1970.13

If a Language has a ts, it has also an s. Greenberg in S 1.3, Hockett UL 25. Follows from the low evolutionary status of the ts: affricates are late and secondary in all languages of the world.

The Consonants kh and h Seldom Occur together in the Phoneme System of the Same Language. This important universal is obscured by the fact that kh and h appear together in the phoneme systems of Classical Greek, Classical Latin (kh = ch < k being a marginal phoneme there), High German, and in the West Slavic languages (h occurs in Polish only on Czech, Ukrainian and Latin loan words). The coexistence of kh (official German spelling ch) and h in High German may be a Slavic substratum (stated by Décsy 1977:1).

Frequency of Occurence

The Number of Principal (Nuclear) Phonemes is Restricted. The phoneme household of a given language may comprise in extreme cases more than 60 phonemes. However, only a relatively small number of them are used as "principal (nuclear) phonemes", i.e. as phonemes which occur in neutral texts with an average more than two percent. Hungarian, for example, has 62 phonemes but only 22 of them qualify as principal phonemes (a ratio of 62/22). The corresponding values for some other languages are: English (Received Pronunciation) 54/20, German 37/17, Sanskrit 48/18, Rumanian 48/17, Lapp 71/18, Votic (a small Finnic language on the south shore of the Gulf of Finland) 75/16, Cheremis 27/16. The poor utilization of the large part of high-number phoneme systems is concomitant with the abundance principle of nature. Backup of this Universal in Décsy 1970.9-11 (with exhaustive statistics from the Finno-Ugric languages).

In Interjections, Marginal Phonemes Occur with High Frequency. Statistical backup for this on the basis of Finno-Ugric languages – Décsy 1970.13. Marginal phonemes occur in general (non-interjectional) texts with a frequency of less than two percent (concept explained in Décsy 1970.9). The low-frequency marginal phonemes in interjections (which are, most often, tone icons) attract attention and fulfill specific emotional duties in communication. The consonant h is a marginal phoneme in both Finnish and Hungarian. However, in interjections, h displays an occurence of 23,6 % in Finnish, and 10,3 % in Hungarian. Statistical details in Décsy1970.13.

The More Different Meanings a Word Has, the Higher its Relative Frequency of Occurence. Ullmann UL 233 after K. G. Zipf. A good example is English get (words with extremely high number of meanings were called "ethnolog" by Wescott, see below)

The Number of Interjections is Higher in Primitive Languages than in Intellectually Developed Languages. Ullmann UL 224. The interjection should be replaced here by descriptive (expressive) words (soundsequences). There are (relatively) new creations of this type even in highly developed languages, cf. English crush, German Krieg, Finnish kapina 'noise, revolution'. See Selkup 329-333 and 152-153 as well as Yurak 55-56.

Homophony/Opacity/Transparence

The Probability of Homophony is Inversely Related to the Number of Principal Phonemes and to the Length of Words. This is, basically, a mathematical banality. For details see Décsy 1970.11.

The Degree of Opacity is Directly Related to the Number of Principal Phonemes and to the Length of the Words. This is a reverse of the previous. See Décsy 1970.12.

The Degree of Opacity is Inversely Related to the Number of Restrictions. Details Décsy 1970.12.

Syllables and Liaison

Syllable is a Universal Phenomenon in the languages of the world, although there is no good theoretical definition of its concept. The syllable is, first of all, an axiom. "The syllable is not only a structural unit but also a necessary unit in speech production, in real speech as well as in models". Georg Heike in K 570. "Syllable is to phonology what phrase is to grammar, namely a unit intermediate between the output and the primes". Saporta UL 66. For syllable concept definitions and difficulties connected with them see Szemerényi 1970.98-99.

In a Language in which all Syllables are Exactly One Phoneme Long, the Distinction between Syllable and Phoneme Disappears. Saporta according to Jakobson UL 267 who regards it as "impossible per definitionem". However, the thesis has great importance for (the presyllabic) Language Origins Times when the "One-Phoneme Unit" was not only a phoneme, but (from a modern point of view) also a syllable, a word, a syntagma, and a sentence (displayed today by some interjections).

The Syllable Boundary Rules may be Set Individually by Certain Languages or Language Groups. Cf. the cases Ba-ntu, Lu-mu-mba in Bantu languages where nt and mb are monophonemic units (the European syllabization would be Ban-tu, Lu-mum-ba), or me-zzo in Italian (Geminata languages like Finnish or Hungarian would syllabize mez-zo, which is unacceptable for Italian).

Liaison is a Natural and Common Phenomenon. Liaison (linkage) realizes in the spoken language the (probably very old) structural principle that every syllable should have a consonant in the initial position (ignored by the script). With this, the spoken language carries on an archaic pronunciation habit. Cf. the English phrase cash and eat (ca-shan-deat RP, or ca-sha-neat in America). The "liason-free" pronunciation of the German Bühnendeutsch is unnatural, new (artificially introduced in the 19th century), and unknown in the dialects. It is a "Lesesprache" phenomenon.

Liaison may Produce Homophony and Ambiguity. Mostly only in oral utterances. The following Hungarian phrases (written differently in each case and having different meanings) are completely homophonous in pronunciation: /hatalmasok/ 'the mighty (persons)', /hat alma sok/ 'six apples are too many', /hatalmas ok/ 'mighty cause', /hatalma sok/ '(he) has too much power'. Native speakers confronted with tape recordings of the four different phrases were unable to identify their original meaning (taped in context).

CV is the most Common Syllable Type in the languages of the world. The types CVC and (especially) VC are late and secondary. Cf. Greenberg's formulation UL.XXV: "Every language has syllables of the form CV (consonant followed by a vowel) in addition to whatever other type it may have". Reich 1986.55 states that syllable types with consonant initial are common in children when they learn the first 25 words (shortly before they turn two). Vowel-initial syllables also occur in the same period, but they do not have chronological priority.

Vocalic Syllable Initial is Unnatural. The type VC and VCV – if they exists at all – always has a kind of glottal stop (harter Einsatz) before the initial vowel. As this pre-initial glottal stop is combinatorially conditioned, most speakers are unaware of the fact that they pronounce it; the official orthographies ignore it. Often an inetymological (prothetic) j, h, v (w), ʔ or ŋ is put before word-initial vowels in order to avoid vocalic beginning, e. g. Czech Vondrák (< Ondrák, petform of Andrew), Yurak nenets 'man' (< enets; in this language, there are no vocalic word beginnings at all). Even estimates based on official scripts (which often disregard initial glottal stops) lead to the view that 80-90 % of the words in the languages of the world have initial consonants.

Initial Consonant Clusters in Syllable and Word Beginnings are Unusual. An estimated 95 % of the languages of the world have no initial clusters. In the CC- syllable initials of Proto-Indo-European, the second C had to be a liquid, and in the CCC- syllable initials the third C had to be a liquid and the second a (separating) stop (p or t), cf. *trem- 'tremble', *dhreugh- 'betray', *strep- 'make a noise', *spreig- 'swell with, abound in'. See Szemerényi 1970.90-91.

In End Position in Syllables or Words, Voiced Consonants (Stops, Fricatives) can be Devoiced/Neutralized. A very common phenomenon. Extremely important in Acrolect German, in the Slavic languages (except for Standard Ukrainian), and in many languages of East Eurasia (Mongolian, Paleosiberian languages). Not known in English or Standard Hungarian (in this language, however, the phenomenon begins to spread, mainly in the Western dialects). In Bühnendeutsch every syllable-closing voiced stop becomes devoiced. Data in Décsy 1973.200 ("Sonorsperre"). A natural precondition of the devoicing is that the voiced consonants occur in final position. In languages like Italian, Japanese, or Finnish (which have mainly words ending in a vowel), the devoicing cannot occur. Nasals and liquids do not devoice, although there are cases when they do (e.g. in Mordvin, partly in Standard French, Spanish). It is remarkable that devoicing is limited to a relatively small number of languages with high speaker number (German, Slavic languages, Mongolian). According to estimates, it may exist in ca. 60-80 languages with a speaker number of about 400 million, mainly in Eurasia.

It is remarkable that English (RP) is not subject to devoicing of the final consonants. Devoicing in American surnames (Bork or Burke < both of German Berg, Grant < French Grand) may display German dialect pronunciations.

Mora

Mora Occurs only in Non-Monosyllabic Tonal Languages. Mora is practically a smaller durational unit within the syllable (mostly of a half-syllable length). E.g. Classical Greek, Lithuanian, Japanese. They may be present in monosyllabic (or monosyllabically oriented) languages too (like Latvian or Chinese), but they are not used for combinatory differentiations there (for which bi- and polysyllabicity is a precondition).

Syllable Tones

24 % of the Languages of the World are Tonal. Africa 70 %, Asia 24 %, North America 19 %, Oceania 15 %, South America 14 %, Europe 5 % (Ruhlen 1977.150). In Africa (where in some areas 90 % of the languages are tonal) and in the Americas, as well as (partly) in Asia, tone may be a retention from very old times. In Europe and (partly) in Asia it maybe a reintroduced feature (like h).

A Language may Contrast Up to Five Levels of Tone, but no more. Ian Maddieson S 2.338.

A Larger Number of Tone Levels Occupy a Larger Pitch Range than a Smaller Number (of tone levels). Ian Maddieson S 2.338.

Phonetically, Central Tones are Unmarked, Extreme Tones are Highly Marked. Ian Maddieson S 2.341.

Systems in which High Tones are Marked are more Frequent than Systems in which Low Tones are Marked. Ian Maddieson S 2.342.

Languages which Permit a Sequence of Unlike Tones in a Word or Morpheme also Permit Like Tones in a Word or Morpheme. Ian Maddieson S 2.343.

A Low (Open Vowel) has a Lower Pitch than a High (Closed) Vowel, Other Things Being Equal. Ian Maddieson S 2.354. This means that the pitch character can depend on vowel quality (timbre).

Vowels on Low Tones are Longer than those on High Tones, Other Things Being Equal. Ian Maddieson S 2.355.

Vowels on Rising Tones are Longer than Those on Falling Tones, Other Things Being Equal. Ian Maddieson S 2.355.

Pause

Pause is a Prosodeme in Every Language. The pause can have semantic function, cf. Hungarian (combined with sentence stress)

tudjátok, hogy □dolgozom 'you know that I am working'
tudjátok, □hogy dolgozom 'you know how I do (my) work'.

According to Varga 1987.101. – In some old manuscript in India the pause was marked by varying distances (interspaces) between words (personal communication of K. L. Janert, Cologne/Germany).

Register

Vocal Registers are an Undeniable Phenomenon of Every Normal Voice. Stated by G. E. Arnold (in Kaiser 1959.48). Every voice shows three main registers: chest, middle, and head register. Details and history of the concept by Arnold in Kaiser 1959.49.

Registers are Sex-Specific. There are three registers both at females and males. The female registers are: 1. Soprano (the highest natural human voice, found in some women and young boy); 2. Mezzosoprano (range between soprano and alt); 3. Alt (the first octave above the treble staff); The male registers are: 1. Tenor (the highest natural adult man voice), Barytone (lower than tenor and higher than bass); 3. Bass (singing voice of the lowest range). Details Lindner 1981.83.

Versification

The Original Function of Verse was Memory Enhancing. Verse as a restricted code is easier to memorize than regular speech.

The Farther Away a Verse Is from Regular Speech, the Higher its Mnemonic and Aesthetic Value. Psychologically, the unusual attracts attention and remains more deeply engraved (etched) in the mind. Note Roman Jakobson's view: poetry is an "organized violence committed on ordinary speech" (quoted after Hallander 1985.58). Note also that Latin *versus* (the source of the English word *verse*) was used originally about the plough turning the earth in the furrow; the *versus* was for the Romans a perverted use of the language as against *pro(or)sus* 'rightly turned' (earth in the furrow or speech in the everyday conversation: *oratio prosa* < *oratio proorsa*). Hallander 1985.58 after Lotz.

The Oldest Metric Systems of Man were Isochronic. Their main unit was the segment or line. The segment or line is in oral performance a pre-assigned time unit (also called "bar") which is filled out mostly by the same number of syllables. However, very often the number of the syllables within the segment or line is different (incorrectly called "irregular", see below). The segments/lines with less syllables than the regular syllable number segment/line have to be performed slower, and the segments/lines with more syllables than the regular syllable number segment/line have to be performed faster than the regular segments/lines. The syllable number differences between the segments/lines give an enormous vividness to the verse in the oral performance. The differences in the number of syllables between the segments/lines are not irregularities but perfectly regular means of the art of versification. This type of versification is common in many so-called developing languages of the world (Africa, Asia, Amerindian languages). It is well known in highly developed Uralic languages like Hungarian and Finnish. The system meets the requirement of the binarity principle in versification: regular segments/lines oppose non-regular segments/lines. The best description of the isochronic versification (after Lotz) is to be found in Hallander 1985.

Binarity is the Ruling Principle in Complex Versification. Complex metric systems utilize only binaric base oppositions, although the language system would in principle admit the use of suprabinaric (more than two member set) oppositional series. In Classical Chinese, six to nine phonemic tonal stresses existed; however, only the binaric "even vs. changing" was used. In Classical Greek, with regard to the relationship between syllabic vowel and subsequent consonant groups, numerous various oppositional sets would be possible; however, only "short vs. long" are used. In English, only "heavy vs. light" (also called "strong vs. weak") is employed. Lotz 1976.234. For complex vs. simple metric systems see Hallander 1985.56.

Versification is Restricted Speech. Therefore, it displays similarities with the speech in early stages of language development. Both early humans and the poets of today had/have to work – compared with present-day structures – with a limited code (limited number of signs). The art of the poet is to find ways within the restricted code to impress us more than we can be impressed by a full code speech of our regular interlocutors.

Rhyme is Secondary and came into being, mostly, independently in languages which have them. Its most important source is line (full-line) repetition. The most perfect (and certainly the most boring) rhyme is a fully repeated line (refrain). Because of the monotony of the full-line repetition, later on as a result of versification development only the terminal parts of the line were reiterated (often in somewhat changed form). Thus rhyme and assonance came into being.

Alliteration is a Versificatory Application of a Linguistic Game common even today among children. In verse, it was used originally for memory enhancing: it is easier to remember words which begin with the same sound.

The Complex Versification Systems of the Traditional Ancient Acrolects are Late and Secondary. Late and secondary means, in this case, even systems which are two or three thousand years old (like Classical Greek or Classical Latin, Chinese, and Japanese). The sophisticated/complex metrics of these acrolects are products of high intellectual interest in versification, which spread with the influence of these acrolects over large areas outside the original territory of their homeland. However, they must have had simple (probably isochronic) forerunner systems.

Complex Versifications Contain More Restrictions than Simple Versifications. This is a basic feature of complexity. Therefore, complex metrics are more different from the regular speech than simple (isochronic) metrics. At the same time, the complex systems are closer to the early evolutionary status of speech than the less regulated simple versifications.

Sound Change

Sound Change Susceptibility. The older (or, structurally: less marked) a customary (non-Proto) phoneme, the higher the probability that it will be subject to a sound change. The consonants p, t, k change more often than s, ch, or sh do. This rule does not apply to m (a very old phoneme), which changes mostly only to n (as a rule, chiefly in word end position; there are good attestations for final m > n in Greek, Finnic, Latvian, and High German). On the other hand, there are seldom sound changes leading to m. Neither does this apply to the new h (=hN), the only possible change of which is disappearance (see following).

In Language Origins Times, hP Changed to k or ŋ. This is a shift of the place of articulation from the larynx to the velum.

The Usual Pronunciation of n Before k and g is ŋ. As a remarkable exception, Standard Russian does not exhibit this sound change.

The most Common Change of hN, r, and l is Disappearance. Follows from the Sound Change Susceptibility Universal. These three consonants disappear in syllable and word end position easier than in syllable and word initial position.

Voiced › Voiceless Neutralization Takes Place Most Commonly in Final Position and Never Intervocally. Ferguson UL 55. The best examples are in the Slavic (dup vs. duba), High German (Rat vs. Räder), and Central Siberian Languages (Mongol, Buryat, Paleosiberian). Can only occur if voiced stops appear in final position (this condition is lacking e. g. in Italian, Spanish, Finnish, Japanese, and Chinese). Trubetzkoy's view (in Grundzüge der Phonologie 1939) on voiced/voiceless neutralization in initial position in Mordvin and in initial-final position in Kirgiz is from a historical standpoint unacceptable. The consonants in question were originally voiceless.

The Most Common Change of Consonants in Intervocalic Position is the Voicing. A good evidence is offered by modern American English.

k › kh › h. French chambre (‹ Latin camara), Hungarian had*(a) 'army' (‹ *kanta 'group of people'), English hundred (‹ hundan ‹ *kundan), all through the stage kh. Ferguson UL 55.

k › t but not vice versa Ferguson UL 55. In Modern Norwegian Lapp, the final k became t in the last century, perhaps under Finnish influence: Finnish has no k in final position (in this language, only t, s, l, m, and r can occur at the end of the words; the glottal stop, originally k, appears assimilated in Sandhi, e.g. taide(?) 'art, Kunst' ‹ * taidek, cf. partitive taidetta ‹ * taide?ta ‹ *taidekta).

t › s Finnish kasi and Hungarian kéz 'hand' (‹Proto Uralic *kati; in Finnish before i only). The most common source of s seems to be t.

s › h but never vice versa. Ferguson UL 55. Cf, Finnish hän 'he/she/it' ‹ *san; Greek ἑπτά 'seven' ‹ *septa, cf. Latin septem, Old Indic saptá).

k > j Proto-Finno-Ugric * mene-k-san '(he) should go' (where the k is marker of the imperative) > Hungarian menjen 'id.'. The change went, certainly, through g (voicing) > gh (spirantization) > j (shift of the place of articulation from the velum to the palatum).

w > v Apparently, labiodental articulation is new in the languages of the world. Originally, every v was a w. This is rather well proven for the Uralic languages and the Slavic languages. I quote Robert Austerlitz: "Modern Europe has very few instances of [w] and is predominantly a [v] area. In terms of geography the solution is to introduce an agent, e.g. the Avars or other Iranian groups with a [v]. A more linguistically founded explanation is that a quinquepartitive vowel system became richer and the [w] as an allophone of [u] was moved into the domain of consonants." Nyelvtudományi Közlemények 76(1976).255.

n > ŋ before Velar Consonants. A universal phenomenon; the only striking exception is literary Russian in which, in words like angel 'angel, cherub', anker 'anchor' (very unnaturally), a simple n (and not an ŋ) appears. As ŋ is a more archaic consonant than n, its nasalized pronunciation maybe a retention: the old ŋ did not change before velar consonants to a dental n.

Cases of Consonant Insertion. They often occur on the syllable boundary between nasals/liquidas and non-nasal/liquidas: **zr > zdr** cf. Old Church Slavonic Izdrael (<Izrael). - **ms(h) > mps(h)** Thompson (<Thomson), Hampshire (< Hamshire). - **ml > mpl** Whampler (< Whamler), Latin templum (< *temlom < * tem-lo- 'place reserved or cut out'). - **nl > ntl** or **ndl** Bentley (< Benley), Pendleton (< Penleton); - **nr > ndr** Russian surname Kondratov (< Konradov < German Konrad). - -**mj > mlj** East- and South Slavic zemlja (< * zemja 'earth, land'; traditional name: l-epentheticon). Such stops are pronounced very often even if the official spelling ignores them, e.g. in English chimney (chimpney).

Balance Lacking Sound Change. This a kind of realignment within the phoneme system either according to the place of the articulation or within the frame of markedness/unmarkedness. If in a language only one member of a customary phoneme opposition exists, the extant member of the opposition can take over (fully or partially) the articulatory peculiarities of the lacking member. E. g. in Finnish there is no ʃ, thus in this language the customary opposition /s vs. ʃ/ does not exist. The result is that the pronunciation of the Finnish s is palatalized (it is pronounced as ç = ś). In Finnish there is no reason to keep apart the articulation of s from ʃ. Similarly, in languages which have only voiceless stops (e.g. Finnish, Estonian, Vogul), the existing stops /p, t, k/ are pronounced slightly voiced (as voiceless media), probably because of the absence of tension as opposed to /b, d, g/ under correlative conditions. This universal has been recognized already by Lauri Hakulinen (see Décsy 1970.14).

There are Fulltonic Languages, and there are Reducing Languages. In fulltonic languages the vowel timbre does not change according to stress type: in unstressed syllables, the vowels are pronounced in the same way as in stressed syllables. Most languages of the world are fulltonic. However, some very important languages (High German, English, Russian, most smaller Germanic languages; also French in "non-mot savant"s) are reducing. Only reducing languages have an /ə/.

Vowel Reduction is the First Stage of the Disappearance of the Vowel or the Entire Syllable of the Word. Danish, French and English are the best-known examples of vowel reduction. Danish and French eliminated numerous syllables in pronunciation. Note that vowel reduction in Russian did not lead to disappearance of syllables.

In Fulltonic Languages, Palindromes Sound Almost Identical in Tape Recorder Backplays. E. g., Latin Anna tenet mappam, madidam mappam tenet Anna 'Anna holds a towel, a wet towel Anna holds" or Roma, tibi subito motibus ibit amor ; Hungarian Indul a pap aludni 'the priest goes to bed'. In backplay, the Latin and Hungarian sentences sound much the same (the only difference is that in backplay the sentence intonation is rising because it is falling in the regular forward reading). There are many palindromes in Hungarian, cf. mozi 'movie'/izom 'muscle', kituno vot rokonok orrtovon utik 'relatives hit the excellent son-in-law square on the nose' (which is, of course, a nonsencical utterance set up only as a linguistic game, cf. József Graetzer: Sicc. Budapest: Móra Konyvkiadó 1976, p. 44-45 with a large collection of Hungarian palindromes). Cf. English pot team tip pit meat top or Sums are not set as test on Erasmus. As English is a reducing language, the backplay of English palindromes is different from the forward run. On English palindromes see Espy 1987.182.

Alternations

Alternations Exist in all Languages of the World. With regard to the utilization of alternations for the differentiation of words and forms (i.e., regular changes in derivational and inflectional sets/paradigms), there are only gradual differences between them. It would be important to put together a list of languages (counterexamples) which have no alternations (if there are any).

Alternation can have Morphophonemic Functions. Cf. English foot vs. feet. German Vater vs. Väter, Finnish (dialectal) tuvas 'in the living room' (Standard tuvassa) vs. tupas 'your living room' (Standard tupasi) (both from tupa 'living room').

Alternations Enhance the Contrasts between Soundsequences within the Paradigm. See the opposites English grind vs. ground; or Finnish yksi (< *ykte) 'one' vs. yhden (genitive sing. of the same word) where only the initial sound is identical between the two forms. The soundsequential oppositions grind/ground and yksi/yhden are much more contrastive than grind/*grinded or yksi/*yksin would be.

Vowel Harmony

Vowel Harmony has Many Different Variations. It is used, very inconsistently, for many different combinatory phenomena. Its cause and its structure (conditions) may be completely different even within the same genetic unit. Cheremis vowel harmony has, e.g. an appearance similar to the Hungarian (palato-velar and labo-illabial); however, its conditioning is different. In Cheremis (which has a mobile accent), the vowel of the stressed syllable determines what kind of vowel occurs in the non-first syllable(s). In Hungarian (where the first syllable is stressed), the vowels of the non-first-syllables follow the timbric character of the vowel in the first syllable. However, there are exceptions.

Vowel Harmony is Always Partial with Regard to the Timbric Vowel Features. This means that there are no absolute specific sets (e.g. u - u, o - o, i - i, etc.), but only generic restrictions (back vowel - back vowel, front vowel - front vowel, labial vowel - labial vowel). The timbric character of the vowels may vary according to results of otherwise relevant historical sound changes – cf. Hungarian /vagyok/ 'I am'' (with o in the second syllable) and /vagyunk/ 'we are' (with u in the second syllable).

Vowel Harmony Makes Speech Monotonous and Limits Distinctiveness. Therefore, it is often eliminated or simplified in structure. Turkish vowel harmony is less restrictive, and therefore more monotonous, than Hungarian, Finnish, or Cheremis vowel harmony.

A Strongly Restrictive Vowel Harmony is More Archaic than a Lenient Vowel Harmony. It is a remnant of times when the particular language had a very limited vowel system.

In Vowel Harmony, the Word is Structured from the Beginning to the End. That means that the character of the vowel in the first syllable determines the timbric character of the vowel(s) in the non-first syllables. Formulated after Heinrich Wagner, cf. Zeitschrift für Celtische Philologie 29(1964).225-298 and UAJb 42(1970).274-275.

With the Umlaut (Apophony), the Word is Structured from the End to the Beginning. That means that the timbric character of the vowel in the non-first (generally the second) syllable is decisive with regard to the timbric character of the first vowel – see Vater (Sg.) vs. Väter (< fatheri) in German. Formulated after Heinrich Wagner, cf. Zeitschrift für Celtische Philologie 29(1964).225-298 and UAJb 42(1970).274-275.

Vowel Harmony and Umlaut (Apophony) Exclude Each Other. A language which has vowel harmony does not have umlaut (apophony) as a structure feature, and vice versa. Occasional (sporadic) counterexamples ("Gegenbeispiele") constitute no evidence against this universal. In Hungarian the old soundsequence /magyeri/ led on the one side to /Magyar/ (selfdesignation of the Hungarian) and, on the other side, to /Megyer/ 'place name form'. In the same language, the Slavic loan челядь appears both as /család/ 'family' (apophony !) and as /cseléd/ 'indoor servant' (by vowel harmony). This type of bifurcation is, however, only sporadic and not structural.

Languages with Vowel Harmony have Vowel Harmonic Parallel Forms. Between the forms, there are only minor semantic differences, e.g. in Hungarian gomboc/gömböc 'dumpling'; in Finnish tuhma 'naughty'/tyhmä 'stupid'; in Manchu ama 'father'/eme 'mother', alioi 'a pipe of a major key/elioi 'a pipe of a minor key', ganggan 'strong/genggen 'weak'. For the Manchu examples see Ohno Susumu: The Origin of the Japanese Language. Tokyo: Kokusai Bunka Shinkokai 1970, 113.

Synharmony

The h, Glottal Stop, and kh are Often Followed by the Same Vowel which Precedes Them. This combinatory phenomenon is called synharmonism. See Finnish (old) talohon 'into the house' (illative singular of talo 'house') and kuuhun 'into the moon' (illative singular of kuu 'moon'). Yurak vahav 'bed', tokhov 'matter', sihiv 'seven', nykhyta 'strong' (Yurak 15).

Aesthetic Value

Aesthetic Value of Sounds is Relative and Subjective. There are views that the more sonorous a sound, the higher its aesthetic value. The most sonorous vowel is a, the most sonorous consonants are r/l. The least sonorous sounds are i/u among the vowels, and h as well as the glottal stop among the consonants. Nevertheless, the highly sonorous r/l are acoustically regarded as unpleasant, and the low saturation i/u are often praised as very attractive vowels. German and Russian poets wrote poems (lipograms) without r (GLC 93), qualifying this high sonority saturation consonant as "ugly". There are English lipograms without e (Espy 157-159); e is in its reduced form (ə), of course, the least sonorous vowel of the phoneme system.

Aesthetic Value of Soundsequences/Words is Subjective. Evaluations are generally phoneme-based, although (subconsciously) concept irradiation may play a major role in the qualifications. The Finnish Nobel Prize winner Sillanpää regarded aurinko 'sun' and the Hungarian poet D. Kosztolányi the female name Ilona as one of the "most pleasant" words of their mother tongues.

2.2 MORPHOLOGY

There are Two Basic Morphological Structure Types in the World. **1**. Analytic. Its oldest subtype is the amorphic language: all words in a sentence keep their original lexemic meaning, there are no affixes (grammemes), e.g. *man work house inside today afternoon. The newer subtype of the analytic structure type is a language in which affixes (grammemes) already exist; however, the original full-lexemic meaning of the affixes is (still) recognizable: father is a teacher. The word is is here a form of the existence verb; however, it is used already as a grammatical element (auxiliary/depletive) because it has (partly) lost its full-lexemic meaning (and appears only as a grammatical element) being subordinated in the sentence to the predicate (teacher), where it indicates verbal categories (like tense, mood, person) which are not characteristic features of the noun (nouns cannot be conjugated). **2**. Synthetic. This inludes the agglutinative, inflective, and incorporative subtypes: the affixes are fused with the stem so that we cannot recognize them as bearers of original lexemic meanings: Latin Omnes homines beatitudini student 'all human beings strive for happiness'; the endings -es, -tudini and -ent (once perhaps independent words, although this is not proven) are not recognizable as (remnants of any) full-lexemic words.

There are No Languages in the World which would Represent a Pure (Clean) Morphological Structure Type. An amorphic language was, perhaps, the pregrammatical speech of the people in the third decamillennium. In every language of the world, the three morphological structure types occur in a different relationship.

Evolutionary Hierarchy between Morphological Structure Types. Amorphic language is the oldest, analytic is younger, and synthetic appears to be the youngest among the morphological structure types.

Amorphic Language is, from a Formalistic Point of View, Ideal. In such a language, every concept and every grammatical form would be denoted by a different sign (soundsequence). It would be an absolute suppletive language (having only paradigms like go vs. went; to be vs. I am), without any hazards of homophony. However, no humans would be able to master such a language because of the enormous burden it would mean for the memory. Nevertheless, computers (which are digital and non-sound-sequential communicators) apply this type of language, in which misunderstandings (form or meaning confusions) are practically nonexistent. See also Makkai's Martian/Venusian/Tellurian Sign Systems (S 3.405-406).

Synthetic Languages often Become Analytic. Chinese and English were earlier analytic languages. Their recent (present-day) analytic structure is the result of a modern development. Structurally, highly developed languages that have reached the peak of their evolutionary history often return to original (primitive, simple, almost-amorphic) stages. They then start a new cycle, developing to analytic and then synthetic, and later on, perhaps again to analytic languages.

The Operational Rules of Simple Grammar are More Complicated than those of a Complex Grammar. This is a mathematical banality. A language which has a small number of grammatical forms (simple grammar) has to use the same elements more often to express the same number of semantic contents than a language which has a large number of grammatical forms (complex grammar). The form inventories of simple grammar languages are small, but their operational rules may be extremely complicated (cf. English or many pidgins and creoles). With other words: in simple grammar languages, you have to express much more with less; the result is lack of opacity. The forms of a simple grammar language are more often ambiguous than are the forms of a complex grammar language.

Simple Grammar and Phraseology/Idiomaticon are Inversely Related to Each Other. The simpler the grammar, the larger the number of stock phrases (idiomatic expressions, locutions).

Parts of the Speech/Word Classes

The Proto Part of the Speech (Proto-Wortart, Proto-Word-Class) is the Interjection. Originally, probably a simple vocsonant ə/ʔ, it became a simple syllable of the type CV (ʔə), then a (still strongly emotional) sememe (word with conventional meaning). Its repetition in continuous sequence created the pattern for bi- and polysyllabicity (ʔə-ʔə, ʔə-ʔə, ʔə-ʔə-ʔə, ʔə-ʔə-ʔə-ʔə, etc.), still in pre-timbre times.

Parts of the Speech (Word Classes) Represent No Universal Categories. They are onesidedly European-oriented and constitute the invention of the Greek-Latin grammarians of Hellenism. Noun, Verb, Pronoun, Numeral, Adverb, Preposition, Postposition, and Conjunction are basically semantic and not grammatical categories. Evolutionarily, they represent the semantic variations of a Proto Part of the Speech, i.e. of Interjections. The Greek-Latin parts of the speech are mostly not applicable for languages outside the European grammatical and language tradition.

Word

The Word is a Basic Unit of Language. In monosyllabic languages the **syllable** is identical with the word. An "open or tacit reference to the word" as a basic unit was required even by the most formalistic-segmentic schools of modern linguistics (Prague Circle, American formalists). Greenberg S 2.5, Ferguson S 3.1-2.

In the Beginning, All Human Languages were Monosyllabic. Follows from the principle that the simple precedes the complex. A self-evident universal, an axiom which needs no proof. At this stage, there was no difference between syllable and word. See today Chinese, which returned to the original monosyllabicity.

Autogenesis is an Ongoing Form of Word Production. Hermann Paul called this type of word creation "Urschöpfung" (Prinzipien der Sprachgeschichte 183) . According to synaesthetic (organs of sense) classes, there are four types of autogenetic word products: those which describe tactile (e.g. pierce, scabrous, Latin scaber), olfactory (mawkish), visual (harsh), and auditory (sound imitating: crush) experiences. The first three are mostly called expressive or descriptive words (they soundsequentially express/describe experiences made in the area of non-auditory senses: tactile, olfactory, visual; there is a shift from experience to expression between different sensual areas). The audio-based descriptive words are called "sound imitating" (as the experience is audio-based, no shift is necessary with regard to the sensual field). Autogenetic words are often regarded as iconic (toconic). However, their iconicity (causal connection with the concept) may be the result of a posteriori reasonings. Most of these words are relatively new. Generally, they play only a minor role in new word production. Their iconic (toconic) character is lost soon after their creation.

Autogenesis Creates Opaque Words/Singlets. Their distinctive value is very high. They are mostly very short and consist of unusual/attractive sounds or syllable combinations. For autogenesis/Urschöpfung see Paul 1960.174.

The Oldest Words of Protolanguages Came into Being by Autogenesis. Etymologically, we can trace such words as English fish or Finnish pää 'head' 'id.' back to the Proto-Indo-European and Proto-Uralic, respectively, but not further. From this it follows that all words of the protolanguages must originally have been autogenetic creations. For the social background of such partly premeditated word creation see Décsy 1985.269 ("inventionalism").

Gender; Class

Gender Classes Emerged only in Languages with Resolute Sex Orientation. Historically, Afroasiatic and Indo-European seem to be primary and consistent sexist languages. This is today ca. 55 % of the world population. In most of these languages the nouns (pronouns) are grammatically assigned to inflectional groups which are natural-sex-based.

Gender is Secondary. It must have come about at a high developmental stage in which the man/woman division must have played a major role. It may be the result of a rigid matriarchy or patriarchy of the early times (supposedly, in the times of the third decamillennium in the southern-central parts of the Old World). If we consider the enormous influence of social structure on language, a coincidental emergence of gender in different areas is imaginable. However, the social preconditions were given for that, apparently, only in the southern-central parts of the Old World. The absolute lack of gender in Northern Eurasia and in the New World is remarkable (in Siberia, only Ket and the extinct Kot grammatized the division masculine/feminine). The non-Indo-European and non-Afroasiatic languages of these areas express the difference between sexes non-grammatically and limited only to natural sex differences.

Grammatical Gender is an Anthropomorphic Metaphor of natural sex and reflects (reflected) the magic-mythological significance of the relationship between male and female in earlier stages of social development (cf. also special "women's languages in Japanese and Chukchi"). However, as with every high freqency phenomenon, grammatical gender lost its original natural-sex-reference. In Classical Greek and Latin it is nothing more than a kind of "excess baggage" which does not contribute substantially to expression.

Gender Always Intersects with the Category of Number. Greenberg S 3.78. There are some exceptions to this in Australian languages.

A Language with Gender Distinction in the 2nd Person Singular Also Has this Distinction in the 3rd Person. Exception: a few languages in Central Nigeria. Greenberg UL.XXXI.

Grammemes/Affixes/Formants

Originally, All Formants/Grammemes (Suffixes, Prefixes, Markers, etc.) were Independent Words. They had complete (full, regular, unlimited) lexical meanings. High frequency words with important generic meaning (to be, to possess, to go, to give, to take) lose their full vocabulary meaning and become grammatical elements (grammemes). U. Weinreich (LU 181) called the process depletion: the prospective grammemes become isolated semantically and then empty with regard to their original full lexical meaning.

A Language which Has No Depletives is Amorphous. In such a language, only full-lexemes can be used. However, such a language can already express almost everything in the same way as a modern language; it can also have a syntax. Modern verbal based computer guided communications are mostly amorphous. The term "amorph" in the meaning 'lacking morphological elements" was introduced by N. J. Marr (GLC 105). See also Makkai's Martian/Venusian/Tellurian Sign Systems (S 3.405-406).

Every Language has Auxiliaries/Depletives. Auxiliaries are full lexemic words subject to (gradual) depletion, during which process they lose their original lexical meaning and become grammatical elements (grammemes) expressing cases, tenses, moods, or other verbal categories. First they constitute analytic grammatical forms (I shall go). Often they become affixes, especially in fixed word order (S. Steele S 3.38). Languages in the early stage of their development had no depletives; modern computer-guided verbal communication systems work without depletives.

High Frequency Full Lexemes in Consistent Juxtaposition Usually Develop to Formants. In a simple pattern, if we have juxtapositions like /house many/, /food many/, /man many/, /child many/, etc., which are repeated many thousands of times a day by the members of a language community, then the lexeme /many/ becomes a grammatical element (sign, formant, marker, suffix) of plurality. It may be shortened to /*-man/, to /*-mn/, or simply /*-m/, and loses its clear full lexical meaning. This procedure is more common in preliterate times, especially in early phases of language origins, than in later times. However, the emergence of the word /like/ as a derivational suffix in English (or the German Leiche to -lich) prove that the force ist still relevant.

Auxiliaries are Intermediate Categories Between a Full-Lexeme and a Formant. Auxiliaries still keep their original lexical meaning; however, they are also used as formants (grammemes) in constructions in which their lexical meaning is unimportant. See is in the sentence /my uncle is a teacher/: in this sentence is is used only as a grammatical element (copula) indicating for the noun teacher verbal categories like mood, tense, person, and number. We may state that we conjugate the noun teacher with the help of is (cf. my uncle was/has been/would be/ought to be/may be/, etc. From this follows that

The Verb is a Conjugated Noun. This also implies the great evolutionary priority of noun over verb.

The Primary Morphological Combination is the Compound. Probably, the first compounds came into being in the timbre time of language development. This means that originally every derivational and inflectional form was a compound. The proto-compound was a simple juxtaposition. It is, at the same time, also the proto-form of the oldest syntactic contructions.

The Mean Length of Morphemes is Inversely Related to the Number of Phonemes in the Inventory. Sol Saporta LU 63 and 70. Basically a mathematical triviality.

Redundance is a Necessary Feature of Soundsequential Language. "If the system is too efficient, mishearing and misconception... become frequent, and ... the speakers are led or driven to make additional distinctions to maintain clarity". Greenberg UL.XVIII. This feature has high relevance in the morphology.

Number

Reduplication is the Oldest Form of Plural. Its primary semantic function was apparently the intensification and (what is the same) the expression of plurality/duality (non-singularity) or continousness/continuity (duration, repetition, frequentativity). It is surprising how often reduplication is used even in modern languages – cf. Chinese rén 'person' vs. rénren 'every person' (see Wang HC 61), Black English good good 'very good', Japanese yama 'mountain' vs. yamayama 'mountains', Hungarian bemegy 'he/she drops in' vs. be-bemegy 'he/she often drops in'. (Moravcsik S 3.299-334 – a very very important study with many many data from many many languages).

Dual may Have Been the Original Form of Plural. Reduplication means/meant, commonly, a one-time repetition. In principle, Japanese yamayama must originally have meant 'two mountains" (iconicity in morphology ?). If this tenet is acceptable, then plural is evolutionarily nothing but a metaphor of the dual. However, the metaphoric meaning in it eliminated its original pre-metaphoric semantics (as very often happens: White House is today basically a pure place name; its original meaning indicating color is not remembered any more).

Dual is a Definite (Exact) Numeral and Plural is an Indefinite Numeral. Dual denotes exactly two; plural, however, means an indefinite number – a sign of the inability of primitive men to count. The plural does not give, but defers the exact information.

When a Language has a Dual, its Plural means 'More than Two' instead of 'More than One'. Clark after Greenberg S 1.247. One of the self-evident universals.

The Number (Numerus) as a Grammatical Category is Universal. It comprises in the inflectional sets not only nouns (and its subbranches, like pronouns) but also verbs (in languages which distinguish between these two parts of the speech).

Dual is a Relic of Times when People Still Were Unable to Count. The appearance of numerals makes the dual superfluous. The regress of dual during the last two or three millennia in many languages (Slavic, Greek, Romance) may have been caused by the rise of exact counting systems.

Dual Can Have Social Motivation. In one of the Munda-Languages (Koh) "it is considered indecent to speak of a married woman except in the dual:

she is, as it were, not to be imagined as being without her husband". Otto Jespersen:The Philosophy of Grammar. London: Allen and Unwin 1924, p. 194. See B. F. Head in S 3.163.

A Mathematically Exact Counting System Makes the Marking of Plural Unnecessary. Digit-oriented computer languages do not use plural markers. Pidgins/Creoles (which are with regard to form system extremely efficiency-oriented) make perfectly understandable statements, saying, e.g. /*three boy see two ship yesterday/.

Numerals Belong to Language Universals because they "constitute a part of the lexicon of any given language". Z. Salzmann Word 6(1950).78. W. Schmidt stated in 1926 that there are no languages which would not have numerals. See Honti in Journal de la Société Finno-Ougrienne 81(1987).133.

Numerals Constitute a Simple Semantic Group Within the Nouns. If we discard tradition, there is no need to maintain a separate part of speech under the label "numerals". The formal features of the numeral are mostly of small significance and secondary.

Numerals are Relatively New. They came about after the introduction of the Dual and Plural. The existence of an exact numeral system and its consistent use would have made the Dual and Plural unnecessary. Proto-Indo-European had the numerals1-10 and 100 (both cardinals and ordinals). Finno-Ugric had the numerals 1-6 and (probably) 10, and 20 as home-grown soundsequences (both cardinals and ordinals). A new review of the Finno-Ugric cardinals is from L. Honti in Journal de la Société Finno-Ougrienne 81(1987).133-149.

Some Numerals are Metaphors of Body or Body Parts. The Uralic *kakta/kȧkta 'two' may be preprotolinguistically akin to *kätä 'hand'. Ugric *koša 'twenty' has been connected with Proto-Finno-Ugric *kuša 'man, male' (the fingers and toes of a man sum up to twenty: a natural counting unit at many primitive peoples e.g. in North Eurasia).

The Higher the Numeral, the More Marked. See Greenberg S 3.255. Concerns only home-grown numerals (not loan numerals). The numeral ezer '1000' in Hungarian (probably from < Middle Persian hazār 'id.'), for example is less marked than kilencszázkilencvenkilenc '999', built up from home-grown elements in this language.

Personal Pronouns/Category of Person

The Most Archaic Group of Pronouns are the Personal Pronouns. Their specific feature in every language is that they know the category of number in every language of the world (even in languages which do not distinguish between noun and verb).

The Personal Endings Originate in Personal Pronouns. The personal nouns triggered the development of personal endings, and with this the emergence of the verb as a formal part of the speech; they are/were the material/physical source of conjugational endings.

The Inflection of Pronouns Often Differs from the Inflection of Nouns. Stated by Schlachter 1968.273. Follows from the thesis that (some) pronouns are evolutionarily more archaic than nouns. The inflection of the archaic nouns may have been established fairly long before the inflection of the nouns.

Derivation/Word Production

Iconicity is not Crucial in the Earlier Stages of Word Production. The oldest words of humans could not be iconic. The existence of a large number of timbre vowels (a, o, u, o, etc.) is the first conditio sine qua non of iconicity: with sequences ʔə-ʔə, ʔə-ʔə, ʔə-ʔə-ʔə (no timbre differences between the syllables), it is impossible to achieve iconic differentiations (or differentiations of any type). Timbred vowels (conditio sine qua non of iconicity) appeared, as proven, only in the late stages of language development. A second important precondition is the ability of reasoning within the language community members: an icon cannot be understood without the recognition of causal connections between sign and object. As reasoning (a product of rational development) is a very recent achievement of humans, it could not have played any role at the very beginning of language development. GLC 38.

Derivation Precedes Inflection. There are only stage (stadial) differences between the two parts of Morphology (Derivation and Inflection). Inflectional sets are conglomerates of high frequency derivatives. This is an evolutionarily reformulated version of Greenberg's Universal 29 (UL 93), which states that "if a language has inflection, it always has derivation".

Non-Class Nouns are Older than Class-Nouns. Non-Class Nouns are the names and those pronouns which can not have an article before them (like I, you, he, we, they, this, those, etc. see below at the category of "species"). According to traditional views, non-class nouns (pronouns as well as names) have no connotations (class meanings, collective meanings; they always denote individuals).

Suffixes are More Common than Prefixes. We estimate that suffixes have a prevalence (predomination) of ca. 80 % over prefixes in the languages of the world. The raason may be an inherited word order prejudice: what is more important (and, this is almost always the stem or head word) appears first, followed by the (designated) suffixoid (independent word which is subordinate to the first word, destined to deplete and to become a suffix). The personal pronouns used before a verb in Modern, are practically prefixes. Consider also the relationship between given name and surname: the surname as the more important word (head word) is on the first place in Chinese, Japanese, Korean, Hungarian and some smaller European areas (Southern Germany and Finnish, only substandard). This recalls the suffixal system; it is probably more logical than the (apparently Latin inversion based) general European system, which is "prefixal".

Prefixes are Newer than Suffixes. They may represent a shift in the customary word order rules of a given language community. Varietas delectat, and a consistent suffixation can become boring. Typical prefix languages are Greek, Latin, Slavic, Lithuanian, Latvian; Germanic had a limited prefix inventory (be-, ga-), the presentday large number of prefixes in High German is late an maybe the result of Slavic structural influence (the same can be stated about Hungarian), see D 213 and Décsy 1977:1.

Diminutives and Their Derivational Priority

Diminutives are the Oldest Suffixes. They are emotional, therefore, they must precede concrete (e.g. the place name suffixes which denote spatial relationships). This tenet has been elaborated on the basis of the Uralic languages by J. Gyorke in: Die Wortbildungslehre der Uralischen Sprachen I. Tartu 1934. Diminutives have the connotation of 'good, pleasant'.

Augmentatives are Less Common than Diminutives. The augmentatives have the connotation of 'unpleasant, ugly'.

Nicknaming is a Universal and Very Old Form of Derivation. There are no languages without derivational petforms. The degree of subjective-emotional saturation may be an important typological indicator. Among others, Italian, South German, Lithuanian, the Samoyed and the Obugric languages utilize a large number of affective suffixes. Other languages are rather restrictive concerning nicknaming (especially the intellectually elaborate languages like English or Standard High German). Russian is more affective than most major languages, cf. idi 'come here'/ vs. idi-ka 'be so nice and come here' (the imperative receives here a deminutive suffix which decreases the intensity of the imperative.

Paradigms/Inflectional Sets

Languages Favor Regular Paradigms. Stated by Clark in S 1.260. Not valid for high frequency words.

Analogy is a Trend Leading to Regular Paradigm. However, it makes the inflectional sets less opaque (more transparent, less contrastive). Nevertheless, the analogy is a powerful trend even in native speakers. James Kilpatrick listed the following analogical alignments from American newspapers (written by native English speaking journalists) in December 1987 : creeped (for crept), grinded (for ground), could have just swam out (instead of swum), clinged (for clung), beared down (for bore down), slayed (for slew), bended (for bend), binded (for bound), swelled (for swollen), see Sunday Herald-Times, December 6, 1987, p. A13 (Bloooomington/Bedford, Indiana).

Analogy is More Powerful in Foreigners' Speech than in that of Native Speakers. It is a factor in pidginizations and creolizations. Cf. cases like foot/foots, hang/hanged in foreigners' talk.

Inflectional Affixes May Have Been Derivational Affixes. Both go back to independent words; however, the inflectional affixes may have gone through a derivational stage during their phylogenetic development.

High Frequency Words Favor Suppletivism Cf. Latin bonus vs. melior, Czech dobry vs. lepši 'good vs. better', English man vs. people It is a wrong concept in the artificial languages that their creator applied regular paradigms (long sound sequences) for high frequency words.

Suppletivism is a Paradigmatic Case of Opacity. It is the type of opacity which occurs within the same paradigm. In this sense, the set English king (noun) and royal can be regarded as derivational suppletive forms.

Cases/Declension

Case/Declension Makes Sense Only in Languages which Maintain a Difference Between Noun and Verb. This is the case in about 50 % of the languages of the world. The difference between noun and verb may often regress in highly developed languages too, e.g. in English, where such "nounverbs" as sleep (both noun and verb) are almost the rule.

Originally, Only Local Cases Existed in all Languages of the World. Abstract/grammatical cases (like accusative and genitive) further temporal /modal/final cases (if any) are secondary and developed from local cases.

Semantically, the Local Cases are Quite Inexact. They only give information which are linear (two-dimesional) and approximative (see prevarication tenets). There are no tridimensional local cases, or local cases which express combined local/temporal/modal/final relationships. These references have to be expressed by adverb (adverbial constructions) in every language.

The Number of Cases Varies between Two and Twenty-four in the Languages of the World. The most developed case system is (as far as ius known) that ofHungarian. John Lotz listed the following 24 cases for this language: nomintive, accusative, inessive, elative, illative, superessive, delative, sublative, adessive, ablative, allative, locative, terminative, dative, temporal, temporal-iterative, causal-final, instrumental, factive (to become something), sociative, essive, formal, distributive, modalitive (modality) – see J. Lotz: Das Ungarische Sprachsystem. Stockholm: Ungarisches Institut 1939, p. 63 and 67. Most cases are new; Old Hungarian (spoken between 1000 and 1350 A.D.) had only six cases (nominative, accusative, locative, superessive, factive, essive). The highly developed Hungarian case system can serve as quasi- model for global (deep structure, semantic-semasiological) investigations on case systems in the languages of the world. For deep structure local case descriptions see Fillmore in Bach-Harms 1-88 and Décsy 1981.54-59.

Accusative and Partitive Endings Function as Postponed Articles even if they emerged as metaphors of concrete (local) cases. The Finnish (Finnic) partitive ending is clearly of local origin (-ta/-tä, originally a lative case suffix, cf. still today ulkoa < *ulkota 'from outside'). It functions today as a kind of despecifizing postponed article and not as a real case. The somewhat puzzling Indo-European accusative ending -m may have been an enclitic.

The Original Function of every Accusative Ending was to Indicate the Definiteness. Therefore, the accusative is not a proper (concrete) case; it has always been (from its conception on) an abstract case. That means that the accusative ending can not be very old. Therefore, the identity of the accusative ending -m in the Indo-European and Uralic languages (mainly in the Finno-Volgaic group and in the Samoyed branch) can not be offered as an evidence for common origin of the two European phyla.

The Genitive is a Secondary Abstract Case because it expresses an abstract relationship (possession). In many languages, it came into being from a lative case (often from dative).

Originally, the Dative was a Concrete Local Case. It denoted motion (allative). Datives often become genitives as result of abstraction. Or they are interpreted as genetive by grammarians.

The Accusative is the Only Case which May Be of non-Local Origin. As the only recognizable funtion of the accusative is to make the noun definite, its ending may have been enclitic. Counterexample: Latin and many Indo-European languages, in which the accusative has/had (as a retention ?) local functions, cf. Latin in Romam '(in)to Rome', Greek εις την πολιν 'into the city' (from here, later in Turkish Istanbul).

Verbalnouns

The Infinitive is a Late Grammatical Category. There were no infinitives in Proto-Indo-European or Proto-Uralic. The actual infinitives of the sub-branches of these phyla are clearly secondary. Infinitives (and any verbalnoun) make sense only in languages which maintain a difference between verb and noun.

Infinitive Endings Originate in Local Case Endings. They came into being from case endings which indicate(d) motion (lative) or motionlessness (locative). English still keeps this archaic semantics in the forms /to go, to write/, etc. The infinitive endings Latin -se (< *re), Greek -εν m, Slavic -ti, Lithuanian -ti (duoti 'to give'),Sanskrit -ni, etc. can be explained from local (locative) endings. H. Krahe: Indogermanische Sprachwissesnchaft 2.87 (Sammlung Goschen 64, 1959).The Hungarian infinitive ending -ni was also a lative case suffix.

Verbalnouns Reflect the Old Undifferentated Part of the Speech (Word Class) Verb/Noun. They may have come into being as a compensation for the lost unity (undividedness) of verb and noun.

Verbalnouns can be Inflected in Case and Number in Many Languages. They can express verbal categories like past tense, contemporality, pretemporality, and post-temporality. There are languages which have 15-20 different verbalnomina (Finnish, Votyak, Cheremis).

Verbalnouns are Convenient Replacements for Subordinate Clauses. Cf. the differences /when I lived in London, I went to school there/ vs. /living in London I went to school (there)/. A syntactic use of verbalnouns instead of subordinating sentences is an important typlogical index.

Possessive Inflection

Possessive Inflection Occurs only in Agglutinative Languages. It seems to be limited as an areal feature to Northern Eurasia. However, it appears in Indo-European languages as borrowing (Armenian, some Iranian subbranches) and – as an accidental curiosity – in the Italian dialect of Neaples where solem < sole mio 'my sun'. According to estimates, 20 % of the languages of the world utilize possessive inflection. Estonian, which is on its way to becoming to an inflective language, lost its possessive inflection.

Languages with Possessive Inflection have no Genitive. The genitive of possessive inflection languages is an innovation (e.g. in Finnish and the other Balto-Finnic languages), cf. Finnish /hanen talonsa/ 'his-his house' (redundant, hanen: genitive of han 'he' and -nsa possessive ending 3rd person).

Possessive Inflection is a Conjugation of the Noun. However, in its basic form, it only knows the person and number as a grammatical category. It does not express tense or mood. In any case, it is a connecting feature between noun and verb.

Languages which Have Possessive Inflection, in Principle do not Apply Possessive Pronouns in Attributive Function. Basically, possessive inflection and possessive pronouns exclude each other because both express the same. Possessive inflection languages introduced possessive pronouns late, and only in the predicative function. They use personal pronouns in the function of the attributive possessive pronoun. cf. Hungarian /az èn hàzam/ 'my house' (èn 'I') vs. /a hàz az enyèm/ 'the house belongs me' (enyèm 'mine', can be only used as a predicate, like English mine). Finnish uses the possessive pronouns in attributive function too (probably under Indo-European influence); the result is an obvious redundance : /minun taloni/ 'my house' where the possession is marked twice (minun 'my' and -ni 'my').

Verb/Conjugation

Conjugation Appears Only in Languages which Differentiate Between Noun and Verb. This differentiatin is (if extant) everywhere secondary. Noun and verb originate in an undivided proto-part-of-the-speech.

Mood

The Imperative is Older than the Indicative. The imperative is often short and unmarked; therefore, it must be older than the indicative (which is, as a rule, more corpulent than the imperative). The need to transmit commands is probably older than the plain information (in indicative) which is mostly not more than a kind of entertainment. The high chronological priority of the imperative has been stressed also in Indo-European Comparative Studies (cf. Szemerenyi 1970.285).

Indicative is a Product of the Reasoning Age. The "commandless" mood (indicative) to communicate must have come up in a time (certainly late) when people were already intelligent enough to be considerate toward each other and to give their interlocutors a chance to discuss events of common concern. The predominance of "indicativeness" is a phenomenon of a world which already has an elaborate time concept. Command is a request to act, indicative is a suggestion to consider. The primary function of information is the cause to do something (command) and not just to think about something. Of course, in highly developed societies, thinking is an action too.

Imperative and Indicative are Universal Moods. We do not know verb languages in which these moods would not exist. In non-verb languages the various moods are expressed by adverbs.

Imperative and Indicative are Mostly Marked by Separate Ending Sets even in Agglutinative Languages. Even if there is a separate imperative mood. Cf. the following reconstructed sets of Proto-Uralic anta- 'to give':

	Indicative	Imperative
1st pers. sing.	anta-ma	(anta-ka-ma)
2nd pers. sing.	anta-ta	anta-ka
3rd pers. sing,	anta	anta-ka-sana

Secondary Moods

Secondary Moods

Except for Imperative and Indicative, all Moods are Late and Secondary. There are a large number of them in individual languages: conditional, conjunctive, subjunctive, auditive (Selkup 247), debitive (Selkup 246), optative, latentive (Selkup 241), reflexive, obligative, precative, adhortative, probabilitive (Yurak 46-48), prohibitive (negative imperative), causative, dubitative, intentional. Late and secondary means here that the moods of this type were developed post-protolinguistically (postphyletically).

Verbal Aspect as a Grammatical Category is Late and Secondary. In Indo-European, only Slavic (earliest around 500 B.C.) and Greek (the latter only in past tense: imperfect vs. aorist) established this category (Szemerenyi in Glotta 65.7, 1985). Most languages express it full-lexemically, e.g. English to go vs. to walk, Hungarian menni 'to go' vs. járni 'to walk'.

Aktionsart

Aktionsart as a Grammatical Category is Late and Secondary. Most languages express it full-lexemically, e.g. in English /he begins to eat/, /he eats it up/, /he is eating/. The Aktionsart is basically a derivational phenomenon in languages which utilize it on a large scale, e.g., in the Uralic languages where, among others, the following classes have been set up: momentaneous, inchoative, ingressive, terminative, durative, iterative, frequentative, perfective, imperfective, resultative.

Continuous Forms

Continuous Forms Exist only in Languages which Have No Prefixes. E. g. in English /I am reading/ or Finnish /olen lukemassa/ 'id'. Finnish is an absolutely prefix-free language; English has only a small number of inherited (Anglo-Saxon) prefixes (be-) and a somewhat larger number of borrowed prefixes (re-, post-, a-, etc.). Nevertheless, English behaves as a prefix-free language. French, developed from the rigid prefix language Latin, is also absolutely prefix-free (of course, we have to disregard the Latin prefixes in the late "mot savants").

Experience Form/Reference

Experience Form/Reference as a Grammatical Category is Late and Secondary. It is a specific type of the verbal moods which indicates whether the speaker knows the subject he speaks (reports) about by hearsay (indirect experience) or by direct observation (direct experience, eyewitnessedly). The indirect experience form is also called (in different languages) auditive, narrative, (preteritum) indefinitum, unwitnessed, quotative (Whorf in connection with Hopi); in Russian the terms zaglaznoje, zaochnoje, neochividnoje, oposredstvovannoje are used for it. In German-language publications by Finnish scholars it is called "Referat" (Osmo Ikola's preferred term). The phenomenon in question is expressed in many languages full-lexemically, mainly in casual speech (...as I hear..., ...I was told..., Hungarian aszongya 'as he says' < acrolect azt mondja). Apparently, this distinction is of great importance; therefore, in many languages it has been grammaticized (often only in past tense). As a grammatical category, it has been proven in varying scope in the following languages: Old Indic, Bulgarian, Baltic (Lithuanian, Latvian), Armenian, all Turkic languages, Estonian, Livonian, Zyrian, Votjak, Cheremis, all Samoyed languages, in Jukagir, Kalmyk, Mongol, Buryat, Evenki, Even, Negidal, Georgian, San, Svan, and in some Amerindian languages. A striking feature is the phenomenon of the Turkish (miş-forms) and Modern Bulgarian (which borrowed it, apparently, from Turkish). The need for differentiation between eyewitnessed/uneyewitnessed must have a legalistic background; therefore, I think that its grammatization is late. See H. Haarman: Die Erlebnisform als grammatische Kategorie. Wiesbaden 1970. Mainstream language universal research paid practically no attention to this important phenomenon and its grammatizations; see, however, the remarks of U. Weinreich UL 162-163 (with references to R. Jakobson 1957).

Tense and Time

The Time Concept is Linear (One-Dimensional) in Our Imagination. Two- or tridimensionality in time is possible only by a direct combination with space. This is additional evidence for the spatial origin and locality-boundness of time.

Past (Tense) Makes Sense Only in Indicative. Imperative in past tense is nonsense. As we speak much more in past tense than in present tense, a decrease in the use of the imperative moods in modern human communication is understandable. In non-intellectual spheres of activities (in certain low-level jobs, in military service, and in the way of life of peasants), the imperative is used more frequently than in high-lével intellect oriented business. It is another question that the indicative can express imperative in certain styles rather consistently or even conventionally.

Tense Category Can Appear in Nouns Too; Nouns Can Be Conjugated. This means that nouns can be conjugated. Weinreich mentions with reference to Greenberg that "in Tupi.. there is a conjugation of nouns for tensĕ" (UL 156). The predicative inflection of the nouns in the Samoyed languages, in Mordvin, and in many Siberian languages is another proof for that, e.g. Yurak man khañenadm 'I am a hunter', man khañenamz 'I was a hunter' (khañena 'hunter'). The possessive inflection in many Uralic and Altaic languages is a kind of conjugation too, connecting noun and verb paradigmatically.

Past Tense Markers Were Originally Lexemes with the Meaning 'Time'. Supposed on the basis of the Finno-Ugric Past Tense marker *-i- which may be connected with Finno-Ugric *ika time, age' (Hungarian ev 'year', Finnish ika 'age of a person, Estonian iga 'id.', Lapp jakke 'year'. Mordvin ije 'id.' Cheremis ij 'id.'). Decsy 1981.61.

Tense as a Grammatical Category is a Metaphor of Space. All temporal adverbs are or were originally local (spatial) adverbs; cf. in, after, before. One of the most self-evident language universals. Cf. Minkowski's term spacetime (German Raumzeit) as a proto-category of the two undivided concepts (combined: "I-Here-Now").

Time-Deictic Adverbs are Never Less Differentiated than the Tense System. That means that there are no more past tenses than there are distinctions expressed by adverbs like yesterday, or ago...Uriel Weinreich UL156.

Past and Future are Even Level Categories Opposed to Present. Cf. Clark S 1.251, who states that children confuse temporal adverbs like /still/ and /already/ or /yesterday/ and /tomorrow/.

As a Rule, the Present Tense is Unmarked. Present markers must have special historical background.

Simple Future is Always Late and Secondary. The present tense is mostly ambivalent and expresses future too.

Future Tenses are More Marked than Past or Present Tenses. Details see R. Ultan S 3.82.

Future Tenses Evolve from Spatial Categories, but they do not develop into the latter (R. Ultan S 3.83 and3.118). This follows from the evolutionary principle that the simple precedes the complex: space is simple and con-crete, time (tense) is complex and abstract. As the future is a tense (time category), it must be secondary in relation to space.

Future Expresses Uncertainty. R. Ultan 3.118. This is a self-evident universal.

If the Future is Periphrastic, the Auxiliary will Always Be a Present Tense. R. Ultan S 3.116. Makes sense only in languages which know the (verbal) category of tense. In the sentence Peter visits Evelyn tomorrow the word tomorrow functions as a kind of auxiliary (temporal adverb), indicating (due to its lexical meaning) the future. In languages which have no grammatized tense (Chinese), the future (or past) can be expressed only by such temporal adverb(ial)s.

Word Length

In Continuous Speech, Word Length is Irrelevant. Because of the liaison (linkage), the boundaries of words and syllables are fused. Words are real units only in enumerations (where they are one-word sentences) or in written form where visual sequence features (interpunction, syllabization) make them recognizable.

Zipf Effect. High frequency long words tend to shorten, especially if the influence of a regulating acrolect is missing (as in slang, in preliterate stages, or in artificially created terminologies of newly invented techniques). The intensive shortening of long Latin words in spoken French (hospital > hotel, embrace < *imbracchiare) may be results of this trend. Cf. also movie < moving picture (or Hungarian mozi < mozgókép, an American-based premeditated innovation of the playwright Ferenc Molnar). See H. H. Clark and E. V. Clark (S 1.264-266). A result of the "least-effort" principle advocated by K. G. Zipf in 1935 and 1949 (S 1.177 and 264). Cf. also Decsy 1970.16 and (with many impressive data from Finnish) Anttila in UAJb 47(1975).10-19. As synthetic languages, Finnish and Hungarian have long words; nevertheless, in areas uncontrolled by acrolectic influence (as slang), they shorten their long words with unusual vehemence.

Anti-Zipf Effect. Short words tend to become longer. Monosyllabics of the early epochs of sondsequential language became longer by juxtápostions; cf. today English (two words) going on vs. (one word) ongoing.

Languages with Short Words Have More Homonyms than Languages with Long Words. "Hence the relative frequency of homonymy in English and French, as compared to... German or Italian". A self-evident mathematical universal. Ullmann UL 235.

Synonyms

Abundance in Differentiation between Synonyms is Late in the development of language. Ullmann UL 231. Synonym abundance is the result of intellectual sophistication.

Synonyms Come into Being in All Languages on the Basis of the Multiple Reinvestability Principle (MRP), which leads through semantic diversifications without formal extensions at the one-word level to synonyms (Makkai S. 3.416).

Words once Synonymous are Subsequently Differentiated In Various Ways and thus Cease To Be Interchangeable. Cf. the differences between the once synonymous quick, fast, rapid, speedy in English. Michel Breal's Law of Distribution. Cf. also Bloomfield's statement: "there are no actual synonyms". Ullmann stated: "We automatically tend to discriminate between synonyms". UL 230.

Homonyms and Phoricity; Situative Reference/Context

Context and Situation are Omnipresent Semantic Factors. With their help, any misunderstanding can be avoided in speech. They are extremely economic means in communication because they need no positive/material (bodily, corpulent) signs for expression. Basically, they are soundsequence and code free signs. They are not processed in the coding system of Broca's and Wernicke's area of the brain; they represent icon-type signs.

Multiple Reinvestment Principle as an evolutionary force in natural language rests on the phoricity of human perception, and consequently on the phoricity of human speech. Makkai S 3.416.

Phoricity Exists in Every Language. Its main versions are exophora (situational reference) and endophora (textual reference). Endophora has two temporal variations: anaphora (reference to what was said before) and cataphora (reference to what will be said later). Makkai S 3.416 after Halliday and Hasan 1967.

Endophora and Exophora Often Suffice to Eliminate All Irrelevant Senses. After Ullmann UL 234. However, these silent means of expression can work only in situation or context. Their referentiality is situation or context bound.

Languages in which Generic Terms, Polysemy, and Homonymy are Prevalent will be Relatively Context-Bound. Ullmann UL 238. The real meaning of the English ethnolog get, e.g. has to be found out from the context.

Situative Reference (Exophora) is Older than Context. Context (endophora) is a concept of the alphabetic script, therefore late and typically European. It is nothing more than the alphabetic-visual projection of the situative reference (exophora). Its subgroups are anaphora (reference to what was said before) and cataphora (reference to what will be said later). This after Halliday and Hasan (1967) and Makkai S 4.416.

Exophora is the Oldest and Most Economic Sign. In order to be express a fact, it does not necesseraly require material sign (soundsequence, icon, or whatever); it is co-perceived on the basis of the situation. Cf. the Russian ethnolog давай[davaj] 'go, give, move' (formally, imperative 2nd person singular of дать[davaj] 'to give' without formal subject or object). As a sign type, exophora is similar to ellipsis (which is, conventionally, a syntactic concept).

Honorifics/Deferentials

Honorifics are Often Expressed by Inexact Forms. In Europe, the Spanish-type honorific is the 3rd person (used in German, Hungarian, Swedish, and Polish). The French type honorific is the 2nd person plural (used also in English, Russian, and most Slavic languages). It seems to be a principle that reverence cannot be expressed by the natural/exact form (which would be the 2nd person singular). Another (unnatural, inexact) personal form has to be taken in order to express social respect. D 210.

Honorific/Deferentials can be Grammatized if the social structure so requires this. This is the case in Chinese and Japanese. The phenomena are (relatively) new (however, well spread and traditional) in both languages. The best way to deal with the problem of honorifics/deferentials is to study Japanese verbal inflection, which is a thorough grammatization of this deep structure phenomenon. A good introduction to the problem is R.A. Lange: 201 Japanese Verbs. Woodbury N.Y.: Barron's 1971, IX-XII.

Honorifics Reduce the Use of the 2nd Person. In modern English, the original 2nd person (thou) died out (you is the second person plural , the same as French vous). Apparently, the addressee, even if present, is regarded as absent in formal address. In Japanese, this led to an elimination of the use of personal pronouns. In this language, practically, everything is in 3rd person.

Plural is a Natural Metaphor for Greatness/Social Power. After Roger Brown by B. F. Head S 3.159 and 195. This may be the source of the pluralis majestatis, although concrete and rational considerations may have played a role in its emergence too. Kings may have spoken for the court (a group of people) in plural and not for themselves as individuals (single persons). The addressees for such pluralis majestatis were then, logically, the 2nd person plural (again a group of people) and not just one person.

2.3 SYNTAX

There is No Language Without Syntax. Syntax combines a large (or a somewhat larger) number of words with each other according to certain universal or individual rules.

Syntax May Be Older than Morphology. An amorphic language may already have syntax. Today the icons guiding passengers in an airport often constitute syntagmata and sentence-like phrase units. Arrow sign pointing to a meeting point sign conveys the message that "if you lost your way and want to meet your fellow passengers, follow the direction indicated by the arrows, come to the area of the meeting point sign, and wait there; your friends may come to that point soon (like you do)".

Syntax is Newer than Vocabulary. As every syntactic unit consists of the combination of at least two concepts (words), the words (vocabulary, lexicon) must be older than the syntax (or its smallest units, the syntagmata).

Parataxis (Coordination) Precedes Hypotaxis (Subordination). Parataxis is more universal (and therefore more archaic) than hypotaxis. The highly developed hypotaxis of the European languages comes through the Latin and Greek medium from the Semitic languages; the Hebrew Bible language elevated hypotaxis to an obligatory pattern for the languages of the Christian world.

All Messages Are Syntagmatic Unless They Are Very Short and Global. Robert Austerlitz in: Seiler-Brettschneider 1985.70. Austerlitz refers here, apparently, to global interjections and/or their archaic forms like laughter, crying/weeping, coughing, snorking, munching (noisy eating), clicks.

Interjections Were Not Only the Oldest Words, but Also the Oldest Sentences. J. Kelemen called the interjections "(one) word sentences" (see UAJb 45.279-280, 1973).

If a Message Consists of More than One Word, it Already Has a Syntax. In two-word messages the question inevitably arises: wich word is the subject and which word is the predicate? In a virtual primitive language without morphology ("amorphic/martial" language), a sentence like /*father teach/ can have four meanings: 1. The father teaches. 2. The teacher fathers. 3. The father is a teacher. 4. The teacher is a father.

In Incorporating Language Structures, One-Word Sentences (Utterances) can Display Syntax. In such cases, the necessarily present morphological elements have to be interpreted as parts of the sentence. Cf. Hungarian látlak "I see you'; in this "sentence" lát is 'to see" (stem, lexeme), k indicates the subject person (1st person singular), and l points to the object person (you). Similar stes exist in the Northwest Caucasian languages which developed the "polypersonal conjugation": they mark the subject person and the object person on the verb. In languages with predicativity inflection (e.g. Mordvin, the Samoyed and some Paleosiberian languages) an adverb like here means that 'he is here". The 1st and 2nd person have an ending; therefore, the ending-less simple adverb is understood due to the zero-morpheme as '(he is) here'.

The Oldest Syntactic Constructions were Juxtapositions (Paratactic Word Sequences). They were identical with simple enunciative sentences without copula, like /house white/ 'the house is white', interchangeable with /white house/ (the latter is a basic attributive construction).

The Sentence is a Hierarchically Organized Group of Words and Word Groups (Syntagmata) in Every Language. The elements of hierarchic organization are to be found already in the morphology, especially in connection with the compounds: in the compound (construction, phrase). /white table/, the word /white/ is subordinated to the word /table/. Syntagma is a term of the French de Saussurean linguistic school; its English correspondence is (mostly) phrase, and its German counterpart is Wortgruppe.

The Number of Syntactic Patterns is Limited. These patterns were analyzed by Western grammarians in the Middle Ages on the basis of the Greek-Latin linguistic traditions. Modern syntactic theories are mostly variations of these traditional scholastic-grammatical views.

The Basic Unit of Syntax is the Sentence. The sentence states that (**1**) who (=subject) (**2**) did (predicate, erroneously called also verb) (**3**) what (object) (**4**.1) where (local adverbial) (**4**.2) when (temporal adverbial) (**4**.3) how (modal adverbial) (**4**.4) why (causal adverbial) (**4**.5) with what purpose (final adverbial, purposiveness). There is no place in this system for the attribute and apposition, which are subordinated parts of the nouns (of any noun) occuring in the sentence (as a part of the sentence).

All Languages Use Ellipsis. Weinreich UL 176. The Roman beggar asking for bread probably said Panem! (morphologically an accusative, syntactically an object). The nominal form includes the predicate 'give' (situatively), and, apparently, also the subject 'you', and the local (directional) adverbial 'to us'. The ellipsis is a syntactic realization of the phoricity.

Predicate and Subject

The Predicate is the Primary Main Part of the Sentence. There are sentences without a subject (cf. Hungarian látlak '[I] see [you]'), but never without a predicate. There are languages in which a simple lexical unit is already a sentence indicating existence (e.g. Mordvin and Samoyed languages – see above).

Subject and Predicate are the Main Parts of the Sentence, although the evolutionary and structural-semantical priority of the predicate is obvious.

The 1st and 2nd Person of the Personal Pronoun as Subjects of the Sentence are Irreplaceable. Cf. /I am going/, /you are going/; in these sentences only I and you can occur as subject.

The Personal Pronoun 3rd Person as Subject of the Sentence is Replaceable with Every Noun. Cf. /he is going/ –> /Peter is going/ –> /the student is going/ –> my uncle is going.

Object

Object

Evolutionarily, the Object is Newer than the Predicate and Subject. The establishment of a connection between two things (nouns) which influence each other as a result of an activity (and its description in a statement or a sentence) is probably one of the important milestones in the development of human intellect.

The Object is a Subordinate Part of the Predicate in the Sentence. A connection between subject and object is only possible through the mediation of the predicate.

Adverbials

The Oldest Adverbials are the Local Adverbials. This follows from the general priority of space over time and modalities. The preferential word order status of the local adverbial in the English sentence (even if late) may be a retention of this very old hierarchic principle.

Temporal Adverbials are Metaphors of the Local Adverbials. Compare /in the house/ vs. /in two weeks/. This is a self-evident universal. See Eve C. Clark's formulation: "language is built on a spatial metaphor" (S 4.114).

There Are a Large Number of Modal Adverbials. An estimated 40 to 50 different subbranches of the modal adverbials can be supposed for every developed European language. Their classification is rather introverted (home-system oriented) in every language, but their basic syntactic patterns are identical; their classifications could be terminologically synchronized.

Modality is a Metaphor of Temporality (Tense). Clark S 1.250 offers, with reference to Greenberg, the sentences /John left after Mary insulted Bill/ (temporal) vs. /John left because Mary insulted Bill/ (causal).

Causal Adverbials are Late in Every Language. They constitute the result of the reasoning abilities of intellectually developed or educated modern humans. Reasoning (finding the cause of a certain phenomenon) is unknown in the animal kingdoom or in primitive human societies. That the event which follows a previous event (a temporal sequence) may be (but is not necessarily) caused by this previous event is an enormous achievement in the development of human intellect.

Finality (Purposivity) is the Newest of the Modal Adverbial Categories. Finality/purposivity is a calculated future-oriented projection of causality. It makes it possible to cause (produce) things on a calculatory basis. Most languages express the finality with a lative (i.e. spatial/motional) construction: e.g. /we are here to enjoy ourselves/.

Word Order

There Is No Such Thing as Free Word Order of the Parts of the Sentence. Every Word Order change in the sequence of the parts of the sentence changes the meaning of the statement (at least slightly). Therefore, word order universals or typologies with regard to sequence of the parts of the sentence (a favorite topic of J. Greenberg, sf. SOV SVO) are highly unreliable.

What Belongs Together Mentally, Is Placed Together Syntactically. Clarks S 1.255. This suggests that spatial arrangement is, in a certain sense, an iconic feature in word order (a favorite idea of H. Seiler and R. Posner in Germany).

There is a Fixed/Rigid Word Order within the Syntagmata. In attributive constructions of some languages the attribute precedes the headword (regens post rectum, see Dezso HS 149); such (preattributive) languages are among others English, German, most Slavic languages (except for Polish, and, partly Slovak), and the Uralic languages. In other languages (e.g. Latin, French, Polish), the attribute follows the headword as a rule. Deviation from this rule (in postattributive languages) attracts attention and is used as a form of emphasis. Cf. the differences in Latin between /frater meus 'my brother'/ vs. /meus frater, (non tuus) '(he is) my bother, not yours/. The first is a neutral, the second an emphatic statement. Other cases of endosyntagmatic fixed word order are: preposition before the headword, postposition after the headword. In Turkic, and in many Siberian languages, the predicate (which can be a verb, but also a noun in cases like /he is a teacher/) has to stand at the very end of the sentence. This is one of the valid absolute positional word order rules in certain language areals.

The Order of the Sequence of the Parts of the Sentence is Not Iconic. There is a strong inclination among representatives of language universal research (especially in the Cologne Project of H. Seiler) to assume that the physical distance of the words from each other in the sentence (a visual-iconic feature) may display cross-linguistic universal semantic values. This may be valid within the syntagmata (see /white round table/ vs. /round white table/ below); however, it is not convincing with regard to the sequence of the parts of the sentence (subject, object, predicate, etc.) within the sentence. In particular, the view may be faulty that the preceding ("forerunner") part of the sentence is more important than succeeding parts of the sentence. Apparently, every language or language group (areal or genetic) develops its own rules with regard to privileged and "disenfranchised" word order positions in the sentence (this statement itself can be regarded as a language universal). In Turkic, for instance, the end of the sentence is the privileged word order position because the predicate can occur only in this place (i.e. at the end of the sentence). In this language thus the end ("runnerup") is more important than the beginning "forerunner". Also in the very common attributive construction /white house/ the second word (head word, noun) is more important than the first word (adjective). To this cf. the concepts focus/presupposition of the American formalism (AG 48).

The Place Immediately before the Predicate is Privileged. Those parts of the sentence which should be stressed (emphasized) appear immediately before the predicate if the general word order rules of the oarticular language permit such a positioning.

Used Car Loan. We can state that in this construction, /used car/ is in its entirety the (subordinated) attribute and /loan/ is the headword. This is clear on the basis of the context. The compound /used car/ itself is already an attributive phrase consistung of /used: an attribute/ and of /car: a noun/. Every different segmenting of the syntagma is a fiction of study desk linguists. The "car loans" can not be qualified "used" (this would be a ridiculous interpretation); "loan" is alone the head word.

White Round Table vs. Round White Table. The relationship between the headword /table/ and the two preceding attributes has been a favorite topic of the Cologne Project on Language Universals (Kölner Universalienprojekt) as well as of the West Berlin semioticist Roland Posner. See Seiler in HS 307. We can speculate that the greater the distance of the attributes from the headword, the more limited (smaller) the quality expressed by the particular adjective. The distance displays here thus a kind if iconicity.

VO Order promotes the Establisment of Prepositions. Hetzron K 177 with reference to Wilfried Lehmann's Proto-Indo-European Syntax. Austin: University of Texas Press 1974, pp. 197, 212, 234.

The Copula is a Late Innovation. Most languages of the world do not use a copula in connection with the nominal predicate. In Chinese, e.g., he [is a] teach[er] is the regular form, meaning from a European point of view ambiguously both 'he is a teacher' and 'he teaches'. See Weinreich UL 177. A copula makes sense only in languages which distinguish noun and verb. See ambiguously both 'he is a teacher' and 'he teaches'. See Weinreich UL 177. A copula makes sense only in languages which distinguish noun and verb. See Helga Harries-Delisle's statement" "Outside of the Indo-European family we rarely find copulas that behave like true verbs" (S 4.438).

Context and Situation are Important Factors in Syntactic Sense Processing. Both context (contextivity) and situation (situativity) are mainly of a visual character, although the factor situativity may be based on memory bits (brain storage units) gained by auditory experience too. They are the final touch in the process which eradicates the last chance of ambiguity in a well functioning message transmission. See the article Kontext and Situation by W. Schlachter in Lakó-Festschrift (Budapest 1981, ed. by G. Bereczki and J. Molnár), pp. 167-178).

Sentence Tone; Interrogatives

Structurally, There Are No Differences Between Syllable Tone and Sentence Tone. Sentence tone is a durationally strongly extended form of syllable tone. If a sentence tone is used in a one-syllable sentence, the acoustic impressions will be identical with those of syllable tone, cf.

English no (enunciative) vs. no (astonishment) vs. no (interrogative)

Speech Addressed to Infants is Strongly Prosodic in Non-Tone Languages. Ferguson S 1.208. "Baby talk tone of voice is universal" (S 1.208). "Infants even a few days after birth can discriminate pitch (tone) differences; ... they pay more attention to high pitches than to low ones" (S 1.208). In my opinion, stress/tone/length and ə/ʔ are used by babies in the first six or so months of their life for signals. They learn real (timbred) vowels and consonants only later. It seems that speech addressed to infants in tone-languages is strongly monotonous (cf. Footnote Ferguson S 1. 208 concerning some Amerindian languages).

The Entscheidungsfrage (Yes-Non Question, Yn-Question) is Older than the Ergänzungsfrage (Non-Yes-Non Question, Wh-Question). Even if the affirmative particles (used as the only possible answer forms for Entscheidungsfrage) are (relatively) new (for the reasons see above). In the Entscheidungsfrage, the question and not the answer contains the decisive (focused) information affirmed or denied by the interlocutor. In the Ergänzungsfrage it is vice versa: the question aims at comprehensive information which must be given by the interlocutor in his own terms. The Ergänzungsfrage is an achievement of modern intensive communication in which the information level of the interlocutors is at approximately the same level. GLC 129.

Interrogation is Often Expressed by Unusual Sentence Tone. The sentence tone often occurs also in languages which do not use (variable) syllabic tone. In English and German (both non-tone languages), the tone of the enunciative sentence is falling, the tone of the interrogative sentence (at least in the Entscheidungsfrage or yes-no sentences) is rising. In Hungarian (where the tone of the enunciative sentence is similar to that in English), the yes-no interrogative sentences have a slowly rising tone which reaches the peak at the penultima of the last word (next to the last syllable), where it then falls sharply. Note that the Ergänzungfrage sentences do not have specific non-enunciative tone in English either, or German, or Hungarian.

Languages with Question Particle or Word Order Inversion in Interrogation Do Not Use Specific Interrogative Sentence Intonation. Finnish always has a question particle in interrogative sentences (in yes-no questions): oli-ko Pekka täällä ? 'Was Peter here ?'. If the question character of the sentence is indicated by morphological means, there is no need for a specific interrogative intonation. The interrogative sentence thus has the same intonation as an enunciative sentence.

Interrogative Pronouns Function as Question Particles. Even in English (of course, only in non-yes-no sentences). /Who was here?/; /When do we leave?/. A specific interrogative intonation in such cases would be overmarking (leading to redundancy). The interrogative sentence has (or should have) in such cases the same intonation as an enunciative sentence. See Ultan S 4.219 (for good demonstrations with Russian interrogative sentences).

Given Information is Pronounced with Lower Pitch and Weaker Stress than new information. Wallace L. Chafe 1976, after Thomas Bearth K 129.

Imperative is Pronounced with Higher Pitch and Stronger Stress than indicative, but usually only when command is intended. This universal is based on experience in many languages.

Affirmation/Negation

The Most Common Form of Affirmation Is Repetition of the Word the Question Aims At. Cf. Latin scisne Latina 'do you speak Latin ?'

Answer: Scio 'yes, I know' (verbatim). Or Greek σὺ ουν ἔλεξας ταῦτα 'did you ask this ?'. Answer: ἔλεξα 'Yes, (I did)'. From this it follows that

Affirmation with Particle is Late and Secondary. Evidence: affirmative particles did not exist in proto-languages (like Indo-European and Uralic), neither did they exist in such later languages as Classical Latin or Greek. However, there is a common Germanic affirmative particle (ja). Proto-Slavic had no affirmative article. The Slavic languages do not have a common affirmative particle either (see Russian da, Polish tak, Czech ano, Slovak hej). Arabic naam 'yes' and Hebrew ken 'id.' differ as well.

Particle Negation is Older than Particle Affirmation. In many languages (but by no means in all), the negative particle contains the archaic consonant n (see above). But note that particle negation is newer than the dichotomic full-lexemic negation (type good/bad).

Compound/Multiple Sentences and Conjunctions

The Asyndetic Compound/Multiple Sentence Connection is Older than the Conjunctional. In primitive languages and archaic dialects, speech coordinating sentences are/were used without conjunction (conjunction-free coordination):

Peter came, Joseph went.

In folklore texts of primitive peoples, such sentence connections are the rule (see Pike HS 56). It is often difficult to decide whether such expressions represent a complex/multiple sentence connection or two simple sentences. In oral performance, the length of the pause between them may offer a clue.

Paratactic Conjunctions are Secondary and Acrolectic. In archaic dialect speech and in primitive languages, conjunctions like

and, but (copulative)
either...or, neither...nor (disjunctive)
but , however, thus , nevertheless (adversative)
so, therefore (conclusive)
i.e., (explanative)
in order (final)

are not known; if known, they are loans from the acrolect. However, their use is limited to acrolectically influenced phrases.

Syntactic Subordination is New. It is characteristic for the Afroasiatic (mainly Hebrew) and Eurasian (Indo-European, Uralic) languages. Its pattern was spread with Christianity in Europe and came through the Greek and Latin Bibles (and church languages) to the vernaculars of Europe. The regular pattern for hypotactic phrases in pre-subordination times was the verbal-noun (gerund, infinitive) construction. Note that the hypotactic (subordinating) conjunctions are pronouns (who, after, when, if = German ob, etc.) which were introduced in the vernaculars of Europe under Greek and Latin influence.

Ergativity

Ergativity is not old. The term ergative was created by Adolf Dirr (1867-1930) and used as an oblique-case name in his "Einführung in das Studium der kaukasischen Sprachen" in 1928 (T. Itkonen in: Nachrichten der Akademie der Wissenschaften in Göttingen I. Philologisch-Historische Klasse, Jahrgang 1975, Nr. 7.3). As in the ergative constructions an oblique case is involved, and the oblique cases are phylogenetically secondary, the ergative cannot be primary. In my opinin (see "species" in this book), the ergative construction is one of many ways to express the definiteness of the object. Its formula is: subject in an oblique case (marked), object in nominative (unmarked); can be used, naturally, only if a transitive verb is the predicate (an intransitive verb can have no object, and the presence of an object is a precondition of ergativity); quasi-construction in English:

*from father	makes	(the) cradle
(subject in oblique case)	(trans. verb)	(def. object in nominative)

The definiteness of the object is expressed thus on the subject. Such "distance markings" are common in the Caucasian languages and in many other idioms – among others, in Hungarian, where the definiteness of the object is marked on the verb: látom a házat 'I see the house' (objective conjugation ending: definite object) vs. látok házat 'I see (a) house' (subjective conjugation ending: indefinite object). Russian and later the German and American ergativity research strongly overestimated the age of the ergative construction. The famous habeo-replacement is, with regard to its formula, very close to the ergative construction; the only difference between the two is that the habeo-replacement uses the existence verb (intransitive par excellence !) while the ergative construction applies a transitive verb: Latin /mihi est liber/, Russian /у меня есть книга[u menja jest knigal/, Finnish minulla on kirja/ (all) ' I have a book', or 'I have the book'.

2.4 SEMANTICS

Paleo-Sememes

The Oldest Semantic Concepts are Dichotomic-Oppositional. They can be called "Ancient Proto-Sememes", and may be subhuman. Oppositional concepts like good/bad, I/you, my/your, warm/cold, bright/dark (the latter two can be metaphors of good/bad) may exist in the animal world. They certainly are pre-soundsequential.

The Basic Local (Spatial) Relationships Are Motionlessness vs. Motion. Motionlessness has no dimensions (it is a fixed point in space). Motion has two basic directions: fusion (allativity) and fission (separation, ablativity).

All Words and Formants (Grammatical Forms) of a Language Can Be Derived from the Basic Dynamic Semantic Opposition /Fission vs. Fusion/. This basic concept with obvious local/spatial/visual (conrete) character is an archaic proto-component of semantics (in the semogenensis), as the (vocalic/consonantic = vocsonantic) laryngeals are in the phonogenesis. Evolutionarily, fission probably precedes fusion.

Out of the Three Primary Features of Space, Only Two Are Grammatically Relevant. The three features are:1. Location (static or dynamic): here-there, from-to. 2. Plane or dimension: front-back, up-down, at-in. 3. Shape: round, square, curved. Shape is grammatically irrelevant (it can be expressed only full-lexemically). After E. C. Traugott S 393-394.

The Concept 'Good' Is a Metaphor of Fusion. Fusion is, in basic biological processes, more constructive than Fission. Caressing, huddling, hug, and handshaking are positive and sympathetic motions. Later, in mythologies and religions, fusion was identified with the 'Supreme Good', i.e. with God.

The Concept 'Bad' Is a Metaphor of Fission. Separation is physically and psychologically connected with pain (even in certain phases of biologically necessary procreation). Badness was identified in the mythologies with the devil (evil, wicked).

Every Concept Contains Emotional and Rational Memory Bits. Their proportionality is different according to the evolutionary stage (phase) of a certain language community or of individuals. In less developed communities or in non-adult humans (children), the emotional elements prevail; increase of information expands rationality. Psychological factors (anger, joy, love) may suspend (parts of) rational treatment/behavior rules.

The Oldest Rational Concept is the Undivided 'I-Here-Now'. The second oldest is the 'You-Here-Now'. Cf. the pedolinguistic observation according to which "children generally acquire 'I' before 'you'" (Clark S. 1.255, 1.251). There are languages in which the words 'I', 'here' and 'now' belong together etymologically ; cf. Latin hic 'here" and 'now'; Hungarian én (< *e-mä-nä) 'I', itt (< *e-tä-tä) 'here'.

The 3rd Person as a Grammatical Category Is the Result of An Abstraction. The 1st and 2nd persons are always within the viewing area (present). The 3rd Person is regarded, as a regular case, as absent. The 3rd person is a "non-I" (Nicht-ich – basically a 2nd person) who is beyond the viewing area. Such space-based imagination for the early phases of language development must be regarded as an enormous rational accomplishment. Cf. Hockett's displacement feature in HC 7. See also E. Moravcsik in S 4.353.

All Things in the World are in 3rd Person Except for I/We and Thou/You. A very important exceptionless universal, as far as I know not discussed in earlier research on language universals. The result of this fact is that the forms of the 3rd person have a disproportionately higher absolute frequency (occurence) than the forms of the 1st and 2nd persons.

Black/White Are Often Variations of the Same Stem. See English black and French blanch 'white' (both from Indo-European *bhel-). Cf. also Hungarian fehér 'white' and fekete 'black' (both from an assumed *päkä; from this stem, probably, also fény 'light'. UAJb 51.55-56, 1979). In Hitigama (a Highland language in New Guinea) mola is 'white' and muli is 'black' (Berlin-Kay 1969.47)

Metaphor; Idioms

All Languages have Metaphorically Transferred Meanings. A self-evident universal. Greenberg UL.XXII.

Metaphor is the Oldest and Simplest Way of New Word Production. The Uralic word *pänkä 'head', for example, received the (secondary) meaning 'up(stairs)' (see Hungarian fel, Finnish päälle). Cf. also English head and ahead. Metaphoric word creation does not require the use of formal elements. Therefore, from a grammatical point of view, metaphor is the simplest way of new word production.

The Linguistic Mechanism for Idiom Formation is the Multiple Reinvestment Principle (MRP). It works in all languages of the world in the same way; it is the metaphoric idiom generator. Makkai S 3.401.

All Natural Languages are Subject to the Multiple Reinvestment Principle (MRP) of their minimal meaningful elements in novel senses, regardless of genetic or typological affiliation(Makkai S 3.416). MRP is the ability to make metaphors.

Metaphor Links "situational reference (exophora) to textual reference (endophora) by virtue of recombining already existing linguistic elements in order to describe, assimilate, approximate, or just stand in for a new situation, some of which are already in memory storage, but some of which were missing". Makkai S 3.416.

Idiomaticity Exists in all Languages of the World. Idiomaticity is a new peculiar meaning on the multiword or phrase level. Makkai S 3.416.

Metaphor is the Oldest Source of Ambiguity, which is in its turn the oldest source of confusion in communication. Therefore:

Metaphoric Ambiguity Calls for Formal Precision. Because of this, metaphoric ambiguity is a stimulus for creation of new derivational or grammatical forms.

About 60 Percent of the Word Meanings are Metaphoric in Every Language. We estimate that a highly developed language (like English) has about 500,000 words (lexemic sounsequences). The 500,000 words have at least 1,500,000 meanings (e.g., get has 30 basic verbal meanings, according to the American Heritage Dictionary).

The Older a Word, the More Metaphoric Meanings It Can Exhibit. This is a self-evident universal. Counterexamples (Gegenbeispiele), if any, should be offered. In other words: the metaphor-susceptibility of old words is higher than that of new words.

Only Modern Languages Have Ethnologs. Ethnologs are words with a disproportionately large number of (metaphoric) meanings, e.g. English get, French coup, Finnish pitää 'to hold to keep to maintain, to consider, to employ, to wear, to treat, to like; must, to have, to be compelled to do something, etc.'. The term ethnolog was introduced by Wescott (see Makkai S 3.426). An ethnolog is the result of an acute sign emergency (Zeichennotstand).

Warmth May Be the Metaphor of Bodily Closeness. Hungarian meleg 'warm' can be etymologically connected with mell 'breast' (both male's and female's): warmth may have been produced (increased, conveyed) by pressing another body (preferably infants) or things (or occasionally animals which have higher body temperature than humans) to the breast area. (Lakó-Rédei: A Magyar Szókészlet Finnugor Elemei 430-431 and 418-419.) - Finnish lämpö 'warmth' (< *lämpä-i) 'warmth', Vepsian läm 'warm' (länboi 'fire', lämin 'lightning'), Votic lämmi 'humid warmth', Estonian lämb 'stifling heat" (lämbima 'to suffocate'), Livonian lämato 'to stifle, suffocate', and Mordvinian (Moksha) lämba 'warmth' can be etymologically connected with the stem of Finnish lämmytä 'to bend, to push' (cognate attested only in Vepsian lämbta 'to bend' – see Suomen Kielen Etymologinen Sanakirja 320); its very hypothetical original meaning may have been 'to bow or bend over a person or thing' (producing warmth by closeness of the body).

Species

Reindividualization and Connotation. Originally, in Language Origins times, only individuals were observable and perceivable. With the emergence of the plurality concept, the notion of class (connotation) appeared: the word house meant an individual house, but at the same time (as an additional new concept) also all houses of the world. Thus a need for reindividualization (exclusion of the class meaning of the word – "de-connotationalization") emerged. The means of this reindividualization (re-concretization) are listed below.

Definiteness is a Comprehensive Concept which has many subgroups. One of these subgroups is generic individualization (reindividualization, reconcretization, separation from all other members of the class). Another subgroup is person identification. The Hungarian objective conjugation defines the object person (whether it is in 2nd or 3rd person: I see you or I see him, or I see any other pronominal or non-pronominal object in 3rd person). The Ob-Ugric (Vogul and Ostyak) objective conjugation singles out the object number (whether the number of the persons is many, two, or one). Thus there are basic semantic (and, partly also formal) differences between the Hungarian and the Ob-Ugric objective conjugation. Only the term is identical, not the substance.

Definiteness is Relevant Mainly in Connection with the Object. The subject as an acting individual is in principle always definite. The object is that part of the sentence which requires a clear and consistent marking of the definiteness/indefiniteness (species). The nominal part of the predicate is always definite (in cases with copula like /Peter is a teacher/).

Nine Ways to Express the Species. Scandinavian linguists call the category of definiteness/indefiniteness "species". Other (less suitable oppositional) terms are: concrete/abstract, individual/class(al), total/partial. Evolutionarily, the definite precedes indefinite. Today, there are at least nine ways to express the category of species: **1.** Article: the table/a table. This seems to originate in Afroasiatic (where it is innate); it may have come from Phoenician into Greek, and from there spread with the Bible (Medieval Church Greek, Church Latin) and sermons (strong influence of preachers on the language of the congregations) in Western European vernaculars. **2.** Word Order: Russian olenja ja ubyl 'I killed the deer' vs. ja ubyl olenja 'I killed a deer'. It is a simplification to state that Russian is unable to express the category of species. In Chinese, the object is definite in OV order and indefinite in VO order, see Steele in S 4.596. Cf. also Mongol (Khalkha) chono nokhoi bariv 'the wolf caught a dog' vs. nokhoi chono bariv 'the dog caught a wolf' (L. Bese in DH 124). Cf. also E. V. Clark's statement: "Word order plays an essential role in indicating definiteness" (S 4.91), perhaps even in languages which have other (formal) means for expression of the species. **3.** Simple Accusative; the original function of the accusative may have been, in all languages which have this case, the denotation of the definite object and not just the object. Cf. Turkish Ali bir kitap ariyor 'Ali is asking for a book' vs. Ali kitabi ariyor 'Ali is asking for the book'. The bir in the first sentence is a "despecializing" adverb/particle (DH 125; cf. Finnish partitive below #7). **4.** Attributive Constructions: white house is more definite than house. **5.** Possessive Inflection: my book may be absolutely definite vs. book. This type of specification (making specific) is common, e.g., in the Uralic and Altaic languages within the framework of the so-called possessive inflection. Therefore, they generally do not develop the article (in Hungarian, e.g., the article was only introduced in the 13th-15th century). According to J. v. Farkas and D. Fokos-Fuchs, the primary function of the possessive inflection was specification. This tenet is controversial, because specification is conco-mitant with any possessive relationship (but not vice versa). **6.** By special forms of the verb. These are the objective (definite) conjugation

(Ugric languages, Mordvin, Samoyed languages) and prefixed (perfective) verbs in the predicate. Examples: Hungarian orvosságát tudok 'I know (a kind of) medicine for that (illness)' (indefinite) vs. orvosságát tudom 'I know the medicine for that' (definite); Russian pil vody "drank (some) water' (indefinite) vs. vypil vodu 'I drank (out) the water' (definite). 7. By case ending contrasts, e.g. in Finnish: luin kirjan 'I read the book' (kirjan acc. sg. of kirja 'book') and luin kirjaa 'I read (some parts) of the book' (kirjaa in this instance from < * kirjatak partitive singular). In Finnish the nominative in itself is definite; indefiniteness is expressed by a "despecializing" case (sse the analog Turkis bir under #3 abowe), the partitive, which is of local origin (having originally meant ablative/separative). 8. By ergative; the real function of the ergative construction (intensively discussed in modern linguistics) is to make the object definite (see below). 9. By context without any formal or prosodic element. This is probably the most common form for the expression of definiteness. It is understood in a "simple manner" (by phoricity, see below) when the noun is definite and when it is indefinite.

There Are Nouns Which Are Innately (ab ovo) Definite. In article languages, no article can be used before such words. These are the proper names, the personal pronouns, the demonstrative pronouns, the possessive pronouns (or possessive suffixed nouns in languages which have possessive inflection, e.g. the Uralic languages), possessive constructions with genitive, and some non-partitive numerals. See the incorrect constructions *I see the Peter, *I see the him, *, I see the that, *he sees the my house, *I see the Peter's house. Spoken (often substandard) German forms like Ich sehe den Peter are analogical and redundant. The words listed in this group have no connotations; they can not denote classes, only indivdual persons or units. The non-connotative words are termed "deictic" words (German Zeigwörter or Zeigewörter) as opposed to the connotative concept words (Begriffswörter). See J. Balázs: Funktionswerte der Pronominalität. Budapest: Akadémiai Kiadó 1973.

Individualization: The Final Stage of Graduality in Specification. The attributive construction white house is less specific than my (one and only) house. The latter is an individualization (absolute specification) possible only in highly developed and legally-rationally oriented languages. The individual concept was the basic experience of man at the primary stage of language origins. Class was developed in the later stages of language unfolding. In modern languages class is often more primary than individual. The class-oriented modern languages use specification (as listed above) to return to the individual (one and only) meaning of the concept (individualization). The aboriginal primitive languages operate(d) with only individualistic concepts; modern languages operate with class and individual(ized) concepts. In highly developed languages the class concept is (often) unmmarked, and the individual(ized) concept is marked, cf. table vs. the table. Oppositions like a table and the table represent redundancy.

Inexactness/Prevarication

Prescreening/Prefiltering and Prevarication. Linguistic messages are only partly precise from a scientific point of view. This universal was termed by Hockett (LU 12) "prevarication". In the statement /Peter went to the library yesterday afternoon/ the precise time definitions (when he left, when he arrived, in which part of the library he spent the time, what he did there, etc.) are not given. The message is incomplete from an absolutist-scientific-technical point of view because it does not "tell the whole story". This is, however, a positive feature of the language. Unimportant details should not be included in the message in order to avoid redundancy (inefficiency). The message should be differentiated only as far as is necessary. Cf. the "window" (Greiffilter) in the biological setting of human senses (vision, hearing) which accepts only a part of the waves reaching the ears (16 Hertz to 20,000 Hertz) and the eyes (from 380 to 700 nanometers). GLC 124.

Emotional/Concrete/Abstract

Emotional Precedes Concrete. The concepts of animals, presound-sequential humans, and newborn humans were/are instinctive-emotional. The de-emotionalization (demythicization) of concepts is (phylo- and onto-genetically) a result of human development (socialization, development of reasoning abilities).

Concrete is Never Instinctive/Intuitive. The concept concrete is always the result of reasoning. The "notions" of animals are basically instinctive, even if they seem to be concrete for humans who observe animal life. In this sense, the concrete is a recent human developmental product.

Concrete Precedes Abstract. H. Kronasser regarded Concrete and Abstract as primary categories of equal status. Evolutionarily, at least in the early stages of language development, the concrete always precedes the abstract. This has been proven by the Hungarian semanticist F. Fabricus-Kovács. Cf. L. Hadrovics in Ferenc Fabricius-Kovács Bibliography. Bloomington, Indiana: Eurolingua 1984, p. 19. See also Ullmann UL 242.

Abstract and Emotional/Affective Have Common Features. Both are opposites to the conrete; both are difficult to define and understand by common/ordinary reasoning. Abstract represents, certainly, a more developed/sophisticated concept than concrete; however, it contains pre-conrete elements too.

Plurality is Always Abstract. Everything that is uncountable is abstract. Abstract concepts are late and secondary. Therefore, plurality is late and abstract. See also partitivity (below).

Partitivity is Always Abstract. Partitivity is a breakdown (fraction) of a concrete individual unit into uncountable parts (mass). Partitivity is a negative (fractional) counterpart of plurality, abstract in the same way as plurality (see above).

Some Archaic Etymons

Directional Prepositions Are Derived from Verbs with the Meaning 'to give'. Stated by Hetzron K 179, with reference to English to and German zu, which etymologically may come (like Slavic do 'in') from the Indo-European verbal stem *dō- 'to give'. As 'to give' is a fusional concept, Hetzron's formulation should be changed accordingly (...come from verbs with the meaning 'motion').

'Many/Much' and 'Settlement' May Come from the Same Stem. Cf. the following chart (the Indo-European forms are cognates with each other, i.e. not only between 'city and 'many'). We suppose that 'many' has semantic priority in such cases, i.e. 'settlement' is a metaphor of 'many'.

Old Indic purúḥ 'many/much and pū́r 'city'
Greek πολύς 'much' and πόλις 'city'
Lithuanian pilus 'full, rich' and pilis 'castle, fort'
Hungarian sok 'much' and Sók 'name of a village'.

Beginning and End Are Sometimes Expressed by the Same Word or by the Same Stem. Cf. Vogul ovyl, which is both 'beginning' and 'end'. Cf. also Old Church Slavonic forms iskonu 'beginning' and koneci 'end', and Old Slovenian (Freising Monuments) is-koni do-koni 'from the beginning to the end' (Berneker: Slavisches Etymologisches Wörterbuch 1.560). These semantics are oriented around the observation that the end parts of an object (e.g. a cane or stick) are both "ends", and that only recent convention or expertism name them differently (beginning and end). The concept becomes relevant only in connection with (abstract) time awareness.

Fingers and Toes. Languages frequently refer to toes as fingers of the foot, but never to fingers as toes of the hand. E. S. Andersen S 3.354.

Privileged/Laudative Directional Positions. Top is preferred over bottom, right over left and front over back. E. Andersen S 3.343 states that this attitude is instinctive even in small children. The preferred words are used figuratively in the meaning 'good', their opposites in the meaning 'bad'. In mythologies, God and heaven are high in the sky, hell is located deep under the earth surface. Right means 'correct', left meant in Old English 'diseased'. Front is a synonym for progress, its antonym is backwardness.

All Case Endings and Prepositions are Locative in Origin. Language is built on spatial metaphor. Stated by Hjelmslev in 1935 and by Anderson in 1971. E. V. Clark S 4.114.

Every Comparison with than Implies a Negative Idea. Stated by Jespersen in 1917, see L. R. Horn S 175.

habeo

The Verb habeo Is a Metaphor of Verbs with the Meaning 'to hold, seize, grab'. Striking examples are Slavic iměti (< jęti ~ьmǫ), Greek ἔχειν (primary meaning) 'to hold', Latin habeo (originally) 'to hold' (< Indo-European *ghabh- and ghebh- 'to give, to receive). Stated by Givón, according to Boeder K 213.

The Concept 'habeo' is a Late Legal Abstraction. Most languages have no habeo-verb. It developed in Latin, Greek, and the Germanic and Slavic languages independently , probably as the result of the influence of property law mindedness of the particular societies.

In Languages with a Marked Nominative, the Verb 'habeo' Exists. Stated by Schlachter 1968.285, with reference to Misteli. This implicational universal is apparently based on observations in Indo-European. There seems to be no causal connection between the two phenomena involved .

2.5 VOCABULARY/LEXEMICS

Expertise(ness): A Plethora of Specific Terms. If a concept and its variations are naturally or technically important in a society, we find in its language a large number of highly differentiated opaque lexemes for the denotation of those particular notions. Cf. the differentiated vocabulary of English for hunting (numerous names for various hunting dogs!) or maritime navigation, the many names in Lapp, Samoyed, or Finnish for snow quality (because they live in the arctic area where there is a long winter, and knowing the snow types is important for skiing or driving a reindeer sled). The abundance of opaque words for snow in Lapp was recognized in 1770 by J. Sajnovics (Beweis, daß die Sprache der Ungarn und Lappen dieselbe ist. Wiesbaden: Harrassowitz 1971, Nachwort). See also Ullmann UL 229.

Vocabulary Units (Full-Lexemes) Have Evolutionary Priority Over Grammatical Forms. We are able to express almost everything by full-lexemic units (words) without grammatical forms; but we cannot express anything (or only very very little) using grammatical forms only, without full-lexemic units (words).

Lexical Fields are the Product of Social and Economic Developments. Elementary concept groups are old, modern concepts are new. Andersen S 3.338-339, Footnote 1 and 3. See Sapir's words: "language is a complex inventory of all the ideas, interests, and occupations that take up the attention of the community" (quoted by Ullmann UL 230).

Specific Elaborateness of a Lexical Field Depends on the Social Preconditions of the Given Community. The statement is untenable that generic names are more common than specific names (H. H. and E. V. Clark S 1.237). On the contrary, in elementary concepts, the specific has an old (mostly singlet, unmotivated, underived) name and the expected generic term is missing (or is late and circumscriptional-scientific), e.g.

specific name (singlet)	missing generic term (no singlet)
hand	extremity
man	human
child	variation of human: young
girl	variation of human: young female
horse	hoofed mammal

Words for Elementary Concepts are Old in Every Language. Mostly they originate in the Proto-Language, e.g. names for body parts, words for elementary activities (to sleep, to drink, to live, to die, to go). See the universals on paleo-sememes, and the word retention stability principle by Dolgopolskij/Shevoroshkin (UAJb 61[1989], forthcoming).

Words for Elementary Concepts are Singlets in Every Language. Out of about 1200 words assumed for the Proto-Finno-Ugric language, only about 30 to 35 are derivates. In these proto-language derivates affective suffixes (*-la*/*-lä*, *-ra*/*-rä*, etc.) prevail, e.g. *šiŋärä* 'mouse'.

Color Terms

Every Language has Terms for 'Black' and 'White'. Stated by Berlin-Kay 1969.2. Black is (or was) apparently the comprehensive term for every color which is not white. It means that the speakers of a given language may see the colors in the proper way, but may have no specific term for their designation, because the particular differentiations are unimportant to them.

Color Terms are Relatively New in Every Language. There are languages even today in which only names for black/white exist (Berlin-Kay 1969.2). In such languages, black has the meaning 'non-white'; i.e., black is the name for all colors of the spectrum except for white.

Color Term Acquisition Follows an Evolutionary Pattern. Valid, obviously, only for the evolutionary development of the present-day European color system (today commonly used throughout the entire world). The color terms are added to black/white in the following evolutionary order: **1.** Red added to black and white (three colors). **2.** Either green or yellow added to 1 (black and white), but never both (four colors). **3.** The missing color of stage 2 (either green or yellow) added to 2 (five colors). **4.** Blue added to 3 (six colors). **5.** Brown added to 4 (seven colors). See Berlin-Kay 1969.27, 33-34, 41, 61, 86-87. In folklore texts from the 19th century (e.g. in Vogul and Ostyak), blue, green (and sometimes even yellow) are mixed up (used inconsistently). Classical Greek and Latin had color names which do not segment according to the present-day European spectrum.

If a Language Contains Eight or More Color Terms, Then It Possesses a Term for Purple, Pink, Orange, and Grey, or Some Combination of These. Berlin-Kay 1969. 2-3.

Two Terms for the Same Color are Rare in the Same Language. Exceptions may have historical grounds; cf. Russian sinij and goluboj (both 'blue') and Hungarian piros annd veres (both 'red').

Lexemic Transparence/Lexemic Opacity

Lexical Transparence and Lexical Opacity are elementary instincts of mankind. Ursula Stephany in K 551. Both instinctive elements are products of the developmental stages in which reasoning (finding the cause behind phenomena) has already been established.

Opacity (Etymological Obscurity) is a Feature of High Frequency Forms. Opaque signs do not have the semantic burden of distracting associations. They make possible a more precise denotation (they are "treffsi-cherer") than do (etymologically) translucent words. (Cf. the suppletive sets good/better, I am/you are/ he is.)

Languages with Many Loanwords are Highly Opaque. Cf. English. Ullmann UL 224.

The Designative Efficiency (Precision) of Opaque Words is Higher than That of Transparent Words. The Greek Akron is a more precise name for the Ohio city than "Crown Point" or "High Point" would be. For the average American, Akron is only the name of an Ohio city; it has no further meanings. The other two names (Crown Point or High Point) are associated with numerous other ("disturbing") concepts ("polysemy") which can be disadvantageous in identification of the concepts (i.e. what is meant).

Sound Changes Can Make the Language More Opaque. This may be one of the reasons for sound changes. Historically (etymologically), camera and chambre (‹ Greek kamara 'vault') or hospital and hotel (‹Latin hospitale) are identical. Their present-day English affiliations are opaque. Cf. also English whitsun and white. The trend to opacity may cause sound (or other changes, e.g. the French introduced tete 'head' because the Latin caput (supposed form› *chap) would not have been distinctive enough.

Sound Changes Can Make the Language More Transparent. This is the case, e.g., in connection with billed 'furnished with a bill' and build 'to erect'. The trend to transparence may cause sound changes.

Negation

Full-Lexemic Opaque Negations are More Archaic than (Morphemic) Particle Negations. The opposition good/bad is more archaic and idiomatic than good/ungood. Also, happy/sad is more natural (and therefore older) than happy/unhappy. L. R. Horn S 4.133.

The Most Archaic Way of Negation is Full-Lexemic. Illustrations: good/bad, happy/sad, to be/to lack. Particle Negations (good/ungood/not good, happy/unhappy/not happy) are secondary. From here it follows that

The Negative Particle Is (Pre-Protolinguistically) Not an Archaic Element of Language, even if it has endophyletically an n in many languages (English no, Russian net, Hungarian ne(m), Yurak niś, Korean ani) which suggest an archaic-iconic background. Many languages have a negative or prohibitive particle (or verb) without an n – cf. Finnish ei (negative verb 3rd person sing.), Vogul ati, Turkish hayir, Indonesian tidak, Arabic la, Hebrew lo, Japanese iya, Swahili hapana, Hawaiian 'a'ole or 'a'ohe, Cheremis uke, Mordvin araś, Votyak övöl, Selkup assa, Greek οὐ(κ)/οὐκ/οὐκί (probably from an Interjection).

A Particle Negation is Always Transparent. Cf. to be or not to be vs. to be or to die. Of course, there are always semantic differences between two such sentences.

Synaesthesia

The Oldest Synaesthetic Metaphors are Tactile-Based. Ullmann UL 243, relating to observations made already by Aristotle. Touch is the most common source of synaesthetic metaphors (cf. piercing sound, hard question).

The Preponderance of Synaesthetic Metaphors from the lower (tactile) to the higher senses (olfactory, auditive, visual) is a semantic universal. Ullmann UL 248.

Smell and Taste

There Are No Specific Terms in the Languages of the Word for the Lack of Taste and Smell. We lack in this sensual area terms like 'blind' (visual area) or 'deaf' (audio-area). The reason may be the decreased importance of these "archaic" senses in modern communications. Taste/smell is most developed in humans at birth (apparently a prehuman feature !). With aging, humans' ability to experience these senses declines in a natural way (often by 50-80 %). See Brain 147 (in References).

The Only Qualifiers of Flavor and Smell are 'good/bad'. Although the number of different typical experiences in connection with taste and smell runs into the hundreds, humans have no specific terms for their denotation. Mostly we define the taste and smell experience by the name of the source in which we experience them: the taste of Tokay wine (almond-like), of escalope Viennaise, of bleu cheese, of French perfume, of Turkish coffee. According to certain theories, the primary odors are minty, ethereal, floral, musky and camphoraceous. Brain 179.

All Smells Are Based Evidently on Seven Primary Odors. Brain 176. However, the seven primary odors have no names in our languages. The two olfactory membrans (one in each of the nasal cavity behind the bridge of the nose) has millions of smell (chemo)receptors; each of these ends in 20 threadlike filaments in touch with the nasal cavity.

There Are Only Four Traditional Flavor Names: sweet, bitter, sour, and salty (Brain 180). They have unmotivated (non-metphoric, "singletic") designations in all languages of the world (at least from a descriptive point of view).

Specific Flavors Are Always Designated Metaphorically, mostly using the name of the most typical object exhibiting the particular flavor: Chili soup, Hungarian paprikás csirke 'Paprika Chicken'; important in qualification of foods, perfumes, scents (chloroform, smoke, gas, etc.).

Sweet is a Metaphor of Good – even though sweet, in certain conditions, can be undesirable.

Bitter, Sour, and Salty Are Used as Metaphors for Unpleasant – even though under certain conditions these flavors can be extremely desirable and pleasant.

Word Retention Stability

The Most Stable Words (Concepts/Objects) are, according to Long Range Interphyletic Etymological Comparisons (without any particular order): 1. the personal pronoun 1st person singular and plural (i/we); 2. two/pair; 3. thou/thee/you; 4. who/what; 5. tongue; 6. name; 7. eye; 8. heart; 9. tooth; 10. no/not;11. fingernail/toenail; 12. louse; 13. tear (as in weeping);14. water; 15. dead; further some body parts (e.g. blood, ear), physical activities (drink, eating, sleep, night), some celestial bodies (moon, sun) and some adverbs (e.g. full). There are no verbs among these words. Stated after V.V Shevoroshkin with reference to H. Pedersen and

Dolgopolskij (UAJb 61[1989], forthcoming). See also New York Times November 24,1987, Science Times C 1 and C4 (interview with V. V. Shevoroshkin).

Borrowings: Words and Calques

Every Language has Borrowings. There are no languages in the word which have not contacted other languages.

A Language Will Borrow Non-Nouns Only if it Borrows Nouns. A rather trivial observation based on language contact studies stated as universal by Comrie 1981.203. For the concept of loan translation see Ullmann UL 239.

Loan Translations (Calques) Exist Only in Languages which Have Loan Words. Based on the following.

There is a Hierarchic Ranking in Methods by which Loans are Borrowed. The oldest and/or most common loans are always nouns (primary concepts) denoting innovative notions (e.g. sugar, army, aspirin); these are the "trivial loans". The sophisticated loans are (in increasing order of contact intensity): verbs, adjectives, adverbs, other parts of the speech, interjections, suffixes (first derivational, then inflectional), reshaping of syntax, remodeling of the morphological system, changes in the phonetic system. I disregarded here names, the borowing of which strongly depends on cultural orienation without serious language contact. The languages of the world are full of Hebrew (Semitic) names today, although the donor(s) (Hebrew, Arabic) have never directly influenced most of the particular receiver languages. For a tentative ranking structure of borrowings, see W. Veenker: Die Frage des finnougrischen Substrates in der russischen Sprache. Bloomington, Indiana and The Hague, The Netherlands: Mouton & Co. 1962, p. 2. See also Ullmann in UL 239. Note that loan translations belong to the sophisticated group of borrowings; if there is a loan translation in a language, then this language already displays a large number of simple lexical borrowings (nouns) and less sophisticated borrowings (verbs, adverbs) from the donor language.

Numerals Are Rarely Borrowed. If they are borrowed, the preceding linguistic influence is very strong. An exemption may be one-sided cultural orientation. The Finno-Ugric languages borrowed '10' and '100' from the Indo-Iranian (partly replacing earlier homegrown words); however, there is no evidence of a previous shocking influence of Indo-Iranian on Finno-Ugric.

2.6. OLINGUISTICS

Geolinguistics

The Number of Languages in the World Is Limited. It was about 2,800 in 1985. This number seems to be constant for the entire past of the human race. However, for about 50 years the number has been decreasing. See Décsy: Statistical Report on the Languages of the World as of 1985, Part I, p. 17.

The Average Global Language Density Index is 54. This means that, theoretically, 54 languages are spoken on every one million square mile of the Earth. The continental values are: Africa 94, Asia 58, South America 30, North America 23, Europa 15, and Australia 8. The computation is based on the supposition that on the landface of the Earth is (without the Antarctica which is inhabitable) 52 million square miles (2800 : 52= 53.85). Details GLC 45-46. From here it follows that

The Virtual Precinct of Every Language on the Earth is 18,518 square miles. That is the exclusive territory which (theoretically) should be covered as national homeland habitat by every natural language of the Earth.

The Number of Acrolects in the World is Limited. The number of languages in the world may be around 2,800 (GLC 83-84); accepted and well functioning acrolects, however, number hardly more than 80-100. The most important are English, German, French, Italian, Russian, Spanish, Arabic, Mandarin; about 30 middle size language acrolects in Europe and about 20-40 in territories outside Europe. Influential historic acrolects were (and partly are – even if not spoken by large language communities) Classical Greek, Classical Latin, Hebrew, Sanskrit, Old Persian.

The Higher the Population Density of a Given Territory, the Lower the Number of Languages There. This means that in such areas a predominant powerful language (which becomes an acrolect) absorbs the less developed small languages. Three-fourths of the world population (five billion in 1987) live in four relatively small areas of the Earth: China (in the east along the coast-line), in India, in North Central Europe, and in the Northeastern United States. In all four areas (except, perhaps, for India), the number of native languages is markedly reduced.

The Stronger the Cultural Coherence of a Given Territory, the Lower the Number of Languages There. Good examples are Europe (excluding the Caucasus, about 60 languages), China (more than three million square miles, only about 80 languages), the Asian part of the Soviet Union (more than five million square miles, about 40 languages), Japan (only two languages).

Peoples on Large Islands and Peninsulas Have An Inclination to Make Their Territory Monolingual. Examples: British Isles, Japan, Korea, Scandinavian countries (historically), Italy, Greece, Madagascar, Australia, and New Zealand. Counterexample (Gegenbeispiel): New Guinea. The spread of Japanese and English in the homelands of the two island peoples (in the British Isles and Japan, respectively) has striking similarities.

Standard Language/Acrolect

Every Acrolect Pretends a Unity/Homogeneity which Does Not Exist. The dialects, the regional spoken forms of the Standard Language, and the differences in the speech of individual members of the language community contradict the supposed unity/homogeneity of an acrolect. The unity/homogeneity of the acrolect is only a virtual blueprint. Standard language has never been a fact, but only an ideal, something which should be.

Language Renewals Have Universal Mechanisms. The psychological basis for such renewals is a subjective (or, often, realistic) feeling that the domestic vernacular is intellectually inferior to a (highly respected) prestige language. The prestige language endangers the existence of the mother tongue. Patriotic intellectuals begin to campaign against the foreign prestige language; they try to enrich the home language with dialect words and new (often unusual) word creations. These neologisms are based exclusively on the domestic idiom; they may demonstrate linguistic isolationism, introvertedness, and xenophobia. The use of the foreign prestige language will be limited to a minimum; the patriotic national community applauds. The home language soon becomes in elaborateness equivalent to the prestige language. After a certain time, domestic linguists recognize that the founding fathers of the language renewal went too far in the patriotic effort when they replaced international words with domestic creations, interrupting integrative cross-cultural connections. However, a revision of the language renewal fails. Conclusion: We regret that the previous generation reshaped the language, creating many neologisms, but we should be fortunate that we have the new words which, although often ugly, meet all the demands of modern communication. In final effect, every language renewal invigorates the linguistic and, very often, the political independence of the domestic community. See UAJb. 59(1987).160.

Acrolects are Always the Result of Intellectual Reshaping of the Language. Acrolects are never identical with a certain natural dialect of the language, even if they are based (phonetically or morphologically) on such a dialect.

Huge Acrolects Grew as the Result of Foreign Occupation of Ethnic Territories. The enormous spread of the 10 super-languages of the word (2.615 billion speakers as of 1985, almost 54 % of the world population) was achieved mainly by political/territorial occupation of foreign territories: Mandarin (770 million), English (415 million), Hindi (290 million), Spanish (285 million), Arabic (170 million), Bengali (165 million), Portuguese (160 million), Indonesian (120 million) Japanese (120 million), and Russian (115) million. Data in Statistical Report on the Languages of the World as of 1985, Part I. Bloomington: Eurolingua 1986, p. 21-22.

Religious Fervor Can Play a Major Role in the Expansion of Huge Acrolects. The well-known cases are Arabic, Spanish, and Portuguese; of the extinct languages Sanskrit, Latin and Church Slavonic.

Elaboration of Domains of Reference Generally Is a Factor of Societal Complexity. Armstrong and Katz in G 226. This stage is reached mostly at acrolect levels.

Script and its Decoding

Originally, All Major Writing Systems Were Iconic/Pictorial, and thus soundsequence free. Their soundsequential encoding is the result of thousands of years of development.

Consonants Are More Important in Writing than Vowels, because there are always more consonants than vowels in a given language. Almost every American will know that the vowel-free sequence MNPLS stands for Minneapolis; but nobody would identify the same name complete in IEAOI (which are the vowels of the same soundsequence). Three vowel languages (like Arabic) still use "unvocalized" scripts; the Arabs write only consonants because the small number of vowels (just three: i, u, or a) are easily identifiable in the written consonant sequences. Such scripts are extremely space efficient.

The Written Word is Always More Phonemic than the Speech (Oral Performance). Script naturally disregards most subphonemic elements of speech. Past linguistic investigations which were based on written texts instinctively represented phonological principles.

The Writing and Reading of Alphabetic Texts Is Not a Visually, but an Auditively Based Communication Form. This types of writing/reading require soundsequential encoding/decoding.

Absolute Alphabetic Literarcy Is a Temporary Phenomenon in Human History. It reached its zenith in the century between 1850 and 1950 in Europe. The new, non-soundsequence script-oriented communication system will make alphabetic writing and reading obsolete. Also, the return to the use of the spoken word in telecommunications (telephone, radio, TV, tape-recorded information storage) instead of written forms of communication (letters, books, minutes, etc.) limits the importance of writing/reading.

Dialect

Every Language Has Dialects. A geographical split in natural languages is common and and well-known. We can observe (geographic) dialect diversifications in artificial languages as well if they are spread and used in various countries. Esperanto, for example, has many different pronunciation "standards" (habits, usages) according to the natural language communities of its speakers (German, Chinese, Hungarian, Russian, etc.). All these Esperanto groups use (mostly unknowingly) the articulation basis of their native natural language when they speak Esperanto.

The Number of Language (Geographic) Dialects in the World May Be Around 15,000. Estimated on the basis of the Handbook of Ethnology by J. C. Leyburn; see Décsy Statistical Report on the Languages of the World, Part I, 1985, p. 11.

Macro-Sociolinguistics

Language Is Not Only a Means of Communication, but Also a Social Glue which keeps together peoples speaking the same vernacular as a mother tongue; the attitude of the members of the speech community toward their vernacular is emotionally bound.

The Relationship of the Speakers of a Language Community to Each Other is Strongly Emotional. This factor is the basis of linguistic nationalism, which has politically devastated Europe in particular during the last centuries. It provided a lesson for the young nations of Africa (Third World) in the 20th century.

Language is a Powerful Means of Discrimination – probably even more powerful than the skin color. States and nations (especially in Europe and Asia) were built up mainly along language (linguistic) lines. Out of the 55 states and state-like units in Europe, 48 are linguistically based. Only Andorra, Belgium, Liechtenstein, Monaco, Austria, The Permyak NO (west of the Ural Mountains in Russia), San Marino, Switzerland, and the State of Vatican City are non-linguistically organized political units in Europe. D 159.

The Longer the Continuous Life of a Language in Its Homeland, the More Diversified the Language Geographically. In Europe, the populations of Denmark and Norway probably have has the longest unbroken language tradition. The dialects are strongly diversified in both countries. The Ukraine is more diversified linguistically than Russia (which is basically a young nation). China in the East, India/Pakistan are other examples.

The Younger a Language in a Homeland Area, the More Homogeneous the Speech Community. Good examples are Turkey (which became Turkish speaking after the 10-11th century) and Hungary (in the Carpathian Basin, Magyar since the 10th century). Chinese is less divided with regard to its dialects in Szechuan (where it is an import language) than in the Peking area (which is its oldest homeland). Similarly, English is more homogeneous in North America and Australia than in the British Isles.

Languages of Nomadic Societies are Homogeneous. Linguistic homogeneity depends on the intensity of communication. Nomadic societies have very intensive communication; their language is therefore only slightly diversified. Good examples are Yurak (reindeer nomads on a territory of about 500,000 square miles) or the old Turkic and Mongolic languages of the mounted nomads in the steppe area between the Danube and the Gobi Desert.

The Acrolect of Assimilated Groups Is More Homogeneous than the Acrolect of Original Native Speakers. In the U.S.A. the Western Dialect (whose speakers are mainly assimilated Germans and East Europeans) is (with regard to dialecs) less divided than New England (whose speakers were mainly British groups with English as their mother tongue).

Language Comparison

There Are Two Basic Structure Types. All languages of the world can be assigned to one of the following two structure types: analytic and synthetic. The classical and oldest variation of the analytic language type is amorphic ("martial") language. Amorphic language has no grammar and no auxiliaries/depletives; except for computer languages, it is no longer used. However, this may have been the original (proto)type of every soundsequential language in the third and fourth decamillennia. A modern variation of the analytic language type are those languages which prefer auxiliaries/depletives (like Chinese, English) and have no (or not many) opaque affixes. All other language types belong to the synthetic type. The subbranches of the synthetic type are agglutinative (affixes recognizable), inflective (affixes not recognizable), and incorporative (a large number of affixes packed into a long word).

Agglutinative Type Features. 1. Richness of inflectional suffixes. 2. Limited homonymy and synonymy. 3. Weak differentiation of word classes. 4. Restricted word order. 5. No grammatical gender (no inflectional classes). 6. Verbal nouns instead of subordinate clauses. 7. Developed system of consonants. 8. Word formation by a large number of derivational suffixes.

Synthetic Type Features. 1. Synonymy of inflectional suffixes. 2. Homonymy of inflectional suffixes. 3. Bisememic morphemes (serv-orum). 4. Strong differentiation of word classes. 5. "Free" word order. 6. Gender or noun classificators. 7. Subordinate clauses instead of verbal nouns. 8. Vowel dominance instead of consonants. 9. Word formation by classes of words. After L. Dezső and Vl. Skalička. HS 151.

There Are No Clear-cut Actual Structure Types. Every actual language is a mix of analytic and synthetic structure elements. No living language can be assigned to either of the two main ideal structure types or to any of their subbranches. In structure type qualifications the grade (degree, relationship) of the two structure types (analytic and synthetic) to each other in factual language is the only acceptable means of characterization.

Agglutinative and Inflective Language Types Represent Two Different Developmental Stages. Agglutinative types are always more archaic than synthetic types. The synthetic type develops from the agglutinative type mainly by fusion of the agglutinative/inflectional junctures, which obscure the boundary between root and affix morphemes.

An Inflectional Type is either Older or Developed Faster than a Comparable Agglutinative Type. A good illustration is the contrast between the Inflectional Proto-Indo-European and the agglutinative Proto-Uralic.

Genetic Relationship Consists in a System of Regular Differences and not in Striking Similarities. English house and Hungarian ház 'house' display a striking similartity both in form and meaning. Nevertheless, the two words are not related. The old form of English house was (probably) *husam (unattested); the Hungarian noun goes back to the early (protolinguistic) form *kota (cf. Finnish kota 'shelter, Lapp hut'). The number of such accdental similartities is very high in languages the history of which is ell known. Hence we can conlude that

As a Rule, Striking Similartities Offer no Evidence of Genetic Relationship. This is true in particular for the realtion of most (geographically) distant languages. This signals a very serious problem of the Long Range Language Comparison (LRLC). A large number of striking similarities between (geograpohically) distant languages are mostly "accidentemes" ("tychemes", see above).

500 Years is Probably Sufficient to Seriously Impair Mutual Intelligibility, and 1,000 Years Will Usually Obliterate It Entirely. Ruhlen 1987.7. This observation is based, apparently, on languages which displayed fast(-paced) development (tempo) during the last millennium (English, French, partly German). In my opinion, there is mutual intelligibility between the following old languages and their modern affiliations Church Slavonic (A.D.10th century)/Russian, Early-Proto-Finnic (around A.D. 100)/ Finnish, Old Hungarian (1350)/Modern Hungarian.

Strong Foreign Influence Speeds Up Language Development Tempo. Classical cases are English (strong Celtic, Normanic/Viking, and French influence), French (Celtic and Latin influence), Cheremis (uninterrupted Chuvash influence after the 4th century A.D., Tatar after the 13th century A.D.), Chuvash (Cheremis influence after the 4th century A.D., Tatar after the 13th century A.D.), Lapp (Scandinavian and Finnish influence during the last 2,000 or so years), Bulgarian (Turkish influence 14th-19th century). Counterexample (Gegenbeispiel) is Danish which has never experienced a shocking foreign influence; nevertheless, it has displayed a fast rate of development tempo for the last millennium. D 127.

Introverted Languages Display Slow Development Tempo. These are languages without strong foreign influence. As foreign influence was less common two, three, or four thousand years ago than during the last millennium (1,000 A.D. till today), most of the old languages (Classical Greek, Classical Latin,and even Classical Arabic, Sanskrit, and Japanese) belong to the introverted group with slow development tempo. Chinese and Turkey-Turkish are counterexamples.

For Ordinary Purposes, Languages Serve as Their Own Metalanguages. U. Weinreich UL 162.

The Best Metalanguage is An Amorphic (Full-Lexemic) Language, because it does not prejudice any grammatical structure (analytic, synthetic).

Indo-European is an Inappropriate Metalanguage. Because it is probably the most idiosyncratic language of the world; it displays numerous individual features which are anything but universal. Nevertheless, these features were regarded by earlier research as "universal", sought and "missed" in all other languages of the world (which were then disqualified as underdeveloped not having features of the Indo-European "metalanguage"). Such features are, among others: marked nominative (-s like in Latin dominus 'lord', res 'thing'), relatively large number of cases (obligatory synthetic character), strongly differentiated tense system, compound vowel -o- (Slavic rybolov 'fisherman' consisting of ryba 'fish' and lovitь 'to hunt', in the compound o replaces -a of ryba 'fish'), existence of the verb habeo, and use of the copula. Today we see that the lack of these features (and not their existence) is universal in the languages of the world.

Linguistic Futurology

Diachrony Has a Futurological Aspect. Diachrony has been understood in earlier research as a historical study dealing only with the past of language development. However, it is clear that the future dimension of time is as legitimate as its past-oriented aspect. The importance if this was emphasized already in1973, see D 230-231.

Large Languages Continue to Grow, Small Languages Continue to Disappear. We can expect that every year about 100-150 languages will die out, until the number of languages in the world is reduced to about 1,000. This will happen between A.D. 2,000 and 2,100.

English is Going to Become the Common Second Language of the World. This is already almost a fact. I quote a remarkable ironic gag from the Tonight Show (NBC) from December 9, 1987: In American science fiction films extraterrestrials are interviewed on a regular basis. They all speak perfect English. In global communication the use of any language other than English is unimaginable.

Reasons for the Success Story English. A unique and fortunate combination of factors enormously furthers its popularity. Its Anglo-Saxon origin makes it easily learnable for Germanic-speaking people (about 160 million in Europe), its Roman/Latin component to Romance-speaking people (130 million in Europe, 500 million outside Europe). Its spelling is not soundsequence-oriented, which makes its script similar to Japanese and Chinese (ideographic-iconic) writing systems (900 million people). Due to colonial language tradition, it is the main means of communication in India/Pakistan and numerous other overseas countries (900 million people). It is already the mother tongue of 415 million people in internationally important countries (U.S.A., British Isles, Oceania). It has never been an aggressive occupational language in Europe (unlike German or French), and is readily accepted by all European peoples in West and East. It is the vernacular of the leading world power U.S.A., and as such enjoys a high reputation in the Eastern Bloc. It has been the common language of international trade, business, and communication since the end of the first world war. It has the highest storage value (GLC 107) in science and modern techniques.

Modern Communication Will Become Increasingly Soundsequence Free. Soundsequential language is extremely complicated and difficult (expensive) to use, because it requires sound-based encoding and decoding. It will be replaced by direct concept transmission from brain to brain by ectric channels. The enormous increase of iconic signs (traffic signs, especially in multilingual Europe) during the last 50 years or so is already a step in this direction. Shrinking of the soundsequence in many large languages (English, Chinese, German, Russian) as a result of vowel reduction and consonant elision may also signal such a development. In direct sign transmission, the question arbitrary/causal does no longer arises. Also, the discriminatory social function of language (as the basis for building and maintenance of nations) disappears.

Emotional Components of Speech Decline. Tone/Pitch, stress and quantity (all biologically archaic limbic-centered features of human speech) are almost completely discarded by synthetic speech production. This influences human speech. Also the male and female registers are becoming closer to each other (Vértes UAJb 61[1989], forthcoming).

REFERENCES AND ABBREVIATIONS

AG	Ambrose-Grillet, Jeanne: Glossary of Transformational Grammar. Rowley, Massachusetts: Newbury House Publishers, Inc.: 1978.
Bach-Harms	(Bach, Emmon & Harms, Robert T., eds.): Universals in Linguistic Theory. New York, etc.: Holt, Rinehart and Winston 1968. See Anttila's review UAJb 42.268-271, 1970.
Balla 1984	(Kálman): On the Measurement of the Prosodic Quality of Vowels. Annales Universitatis Scientiarum Budapestiniensis de Rolando Eotvos Nominatae. Sectio Linguistica 15(1984).41-47.
BCD1984	(Butterworth, Brian; Comrie, Bernard; Dahl, Osten; eds): Explanations for Language Universals. Berlin - New York-Amsterdam: Mouton 1984.
Berlin-Kay 1969	(Berlin, Brent & Kay, Paul): Basic Color Terms: Their Universality and Evolution. Berkeley and Los Angeles: University of California Press 1969.
Brain	The Brain: A User's Manual. The Diagram Group. New York: Berkley Books (Edition 1983).
Comrie 1981	(Bernard): Language Universals and Language Typology. Chicago: University of Chicago Press.
Décsy1970	(Gyula): A Search for Universals in the Finno-Ugric Languages. UAJb 42.9-17.

D(écsy 1973)	(Gyula): Die linguistische Struktur Europas. Wiesbaden: Otto Harrassowitz 1973.
Décsy 1977	(Gyula): Sprachherkunftsforschung I. Einleitung und Phonogenese. Wiesbaden: Otto Harrassowitz. Bibliotheca Nostratica 2/1.
Décsy 1977:1	(Gyula): Slavisch-Hochdeutsche Interferenzen. Gerhardt-Festschrift (see GLC 139).
Décsy 1981	(Gyula): Sprachherkunftsforschung II. Semogenese/Paläosemiotik. Bloomington, Indiana: Eurasian Linguistic Association. Bibliotheca Nostratica 2/2.
Décsy 1985	(Gyula): Review of G (=de Grolier 1983). Reviews in Anthropology 12.264-271.
DH	Dezsö, L. & Hajdú, P.: Theoretical Problems of Typology and the Northern Eurasian Languages. Amsterdam: B. R. Grüner 1970.
Espy 1987	(Willard R.): The Game of Words. New York: Bramhall House
G	De Grolier, Eric (Ed.): Glossogenetics. The Origin and Evolution of Language. Chur- London - Paris - Utrecht - New York: Harwood Academic Publishers [1983].
GHb	Gunnemark, Erik & Kendrick, Donald: A Geolinguistic Handbook. Kungalv/Sweden: Goterna 1986.
GLC	Global Linguistic Connections. Bloomington, Indiana: Eurolingua 1983. Bibliotheca Nostratica Vol. 5.
Gregerson	(Kenneth J.): Pharynx Iconism as Figure and Ground. Invited Speaker's Address. LACUS Forum Inaugural Lecture, August 12, 1986, University of Texas/Arlington.
Hallander 1985	(Lars-Gunnar) A Bird's-Eye View of Metrics. In: Wickman 1985.55-71.

Hartshorne1973	(Charles): Born to Sing. An Interpretation and World Survey of Bird Song. Bloomington and London: Indiana University Press.
HC 1982	Human Communication. Language and Its Psychological Bases. Ed. by William S.Y. Wang. Readings from Scientific American. San Francisco: W.H. Freeman and Company.
hN	"New h"; h which came into being as a result of a change from another consonant (e. g. from k, t, s, kh).
hP	"Proto h"; occlusive counterpart of the Glottal Stop in the early stages of language development; protoform of the supralaryngeal consonants.
HS	Seiler, Hansjakob (ed.): Language Universals. Papers from the Conference held at Gummersbach/Cologne, Germany, October 3-8, 1976. Tübingen: Gunter Narr Verlag 1978. Tubinger Beitrage zur Linguistik 111.
Jakobson 1957	(Jakobson, Roman; Huttl-Worth, Gerta; Beebe, John Fred): Paleosiberian Languages. A Bibliographical Guide. New Haven: HRAF Press 1957. Behavior Science Bibliographies.
Jakobson 1979	(Roman & Linda Waugh): The Sound Shape of Language. Bloomington & London: Indiana University Press.
Jazyki	Jazyki narodov SSR. I-V. Moskva-Leningrad: Akademija nauk 1966-1068.
Julliard 1977	(Alphonse): Linguistic Studies Offered to Joseph Greenberg on the Occasion of his Sixtieth Birthday. Saratoga, California: Anma Libri & Co.
K	Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler. Herausgegeben von Gunter Brettschneider und Christian Lehmann. Tübingen: Gunter Narr Verlag 1980.

Kaiser 1957	(L.): Manual of Phonetics. Amsterdam: North-Holland Publishing Company.
Khrakovskij 1977 (V.S., ed.)	Problemy lingvisticheskoj tipologii is tyruktura jazyka. Leningrad: Nauka 1977.
Lass	(NormanJ.): Speech, Language, and Hearing. I. Philadelphia etc.: W. B. Saunders 1982.
Lindner 1981	(Gerhart): Grundlagen und Anwendung der Phonetik. Berlin: Akademie-Verlag.
Lotz 1976	(János) Szonettkoszorú a nyelvrol [Sonetti a Corona on the Language]. Edited by Gyorgy Szepe. Budapest: Gondolat. - A collection of Lotz's articles translated into Hungarian from English. Many of the items were extended and adapted for the Hungarian translations by Lotz shortly before his death; they are the most authentic final versions of Lotz's articles on metrics and some other aspects of linguistics.
LRLC	Long Range Language Comparison
LU	Language Universal
Meier1961	(Friedrich Georg) Das Zero-Problem in der Linguistik. Berlin: Akademie-Verlag. Schriften zur Phonetik, Sprachwissenschaft und Kommunikationsforschung, Nr. 2.
Nauka 1974	Univerzalii i tipologicheskie issledovanija. Meshchaninovskie chtenija. Moskva: Nauka.
Paul 1960	Prinzipien der Sprachgeschichte. Sechste, unveranderte Auflage. Tübingen: Max Niemeyer Verlag.
R	Ruhlen, Merritt: A Guide to the Languages of the World. Stanford: Personal Edition1975.
Reich 1986	(Peter A.): Language Development. Englewood Cliffs, New Jersey: Prentice-Hall (A modern handbook of pedolinguistics enhancing the principle of recapitulationism).

RP	Received Pronunciation. Definition in the introduction of Daniel Jones's English Pronouncing Dictionary.
Ruhlen 1977	(Merritt): The Geographical and Genetic Distribution of Linguistic Features. In: Julliard 1977.137-160. World spread of a number of phonetic features based on a corpus of almost 700 languages.
Ruhlen 1987	A Guide to the World's Languages. Volume 1: Classification. Stanford, California: Stanford University Press.
S	Universals of Human Language. I-IV. Edited by Joseph H. Greenberg. Associate Editors Charles A. Ferguson, Edith A. Moravcsik. Stanford: Stanford University Press 1978. Volume numbers indicated by Arabic numbers.
SB 1985	Hansjakob Seiler & Gunter Brettschneider (eds.): Language Invariants and Mental Operations.International Interdisciplinary Conference held at Gummersbach/ Cologne, Germany September 18-23, 1983. Tübingen: Gunter Narr Verlag 1985.
Schlachter 1968	(Wolfgang): Arbeiten zur strukturbezogenen Grammatik. Munchen: Wilhelm Fink Verlag.
Sebeok 1986	The Problem of the Origin of Language in an Evolutionary Frame. Language Sciences 8(1986).169-176.
Sharadzenidze 1982	(Tinotin Semenova): Tipologija jazykov v sinkhronicheskom i diakhronicheskom plane. Tbilisi: Izdatelstvo Metsniereba 1982; Universals on pp. 85-93.
Selkup	(Kuznecova A. Ji., Khelimskij, Je. A., Grushkina, Je. V.): Ocherki po selkopskomu jazyku. Tazovskij Dialekt. Tom I. Moskva: Izdatelstvo Moskovskogo Universiteta 1980.
Suomi 1985	(Kari): Suomen /h/:n fonologisesta tulkinnasta[On the Phonological Interpretation of the Finnish /h/]. Virittäjä 89(1985). 337-350. Stresses the semivowel character of /h/ - an important feature of this archaic consonant which allows its vocsonantic interpretation /h/a/, see Décsy 1985.266

Szemerenyi 1970	(Oswald): Einführung in die vergleichende Sprachwissenschaft. Darmstadt: Wissenschaftliche Buchgesellschaft.
UAJb	Ural-Altaische Jahrbucher/Ural-Altaic Yearbook. Continuation of the Ungarische Jahrbucher (since 1921). Formerly in Berlin and Wiesbaden/Germany, now in Bloomington, Indiana, published by Eurolingua.
UL	Universals of Languages. Second Edition. Edited by Joseph H. Greenberg. Cambridge, Massachusetts: M.I.T. Press 1966
Uspenskij 1965	(Boris Andrejevich): Strukturnaja tipologija jazykov. Moskva: Nauka. A List of Synchronic Universals from the Area of Phonology and Grammar 182-221.
Varga 1987	(László): Prozodémák a magyar beszédben és jelolésuk az intonációs átiratban[Prosodemes in the Speech of Hungarians and Their Marking in Intonation Transcript]. Mühelymunkák. Budapest: MTA Nyelvtudományi Intézete 3.91-119
Wickman1985	(Bo) Symposium in Memoriam János Lotz (1913-1973). Stockholm: Almqvist & Wiksell International. Acta Universitatis Stockholmiensis. Studia Hungarica Stockholmiensia Nr. 4.
Yurak	Décsy, Gyula: Yurak Chrestomathy. Bloomington, Indiana: Indiana University 1966. Uralic and Altaic Series Vol. 50.

ADVERTISEMENT

Bibliotheca Nostratica ISSN 0342-4871

Monographs on Interphyletic Linguistics, Language Origins Research, Pre-Protolinguistics, Long Range Language Comparison, Global and Intercontinental Linguistics.

Vol. 1 Eurasia Nostratica. Festschrift für Karl Heinrich Menges. Herausgegeben von Gyula Décsy und Christo Dimov-Bogoev. Wiesbaden: Kommissionsverlag Otto Harrassowitz 1977. Band 1. 217 pp. Band II. 256 pp. (in einem Band, broschiert). ISBN 3-447-01834-8. Distributed by Eurolingua. Price: US $86.00.

Vol. 2 Band I. Gyula Décsy: **Sprachherkunftsforschung I.** Einleitung und Phonogenese/Paläophonetik. Wiesbaden: Otto Harrassowitz 1977. 87 pp. ISBN 3-447-01861-5. Distributed by Eurolingua. Price (broschiert): US $22.00. – Band II. Gyula Décsy: **Sprachherkunftsforschung II.** Semogenese/Paläosemiotik. Bloomington, Indiana: Eurasian Linguistic Association 1981, 78 pp. ISBN 0-931922-06-2. Distributed by Eurolingua. Price (broschiert): US $22.00.

Vol. 3 Karl August Wittfogel: **China und die osteurasische Kavallerie-Revolution.** Wiesbaden: Kommissionsverlag Otto Harrassowitz 1978, 69 pp. ISBN 3-447-011924-7. Distributed by Eurolingua. Price (broschiert): US $20.00.

Vol. 4 Wilhelm Pröhle: **Vergleichende Syntax der uralaltaischen (turanischen) Sprachen.** Autograph-Ausgabe. Wiesbaden: Kommissionsverlag Otto Harrassowitz 1978, 229 pp. ISBN 3-447-02003-2. Distributed by Eurolingua. Price (broschiert): US $66.00.

Vol. 5 Global Linguistic Connections. Published on the Occasion of the World Communication Year. Contributions by R. Anttila, Gy. Décsy, C. T. Hodge, J. R. Krueger, J. Lotz, A. Makkai, M. Ruhlen and Th. A. Sebeok. Bloomington, Indiana: Eurolingua 1983, 151 pp. Hard Cover. ISBN 0-931922-12-7. Price: US $26.00.

Vol. 6 Statistical Report on the Languages of the World as of 1985. Compiled by Gyula Décsy. Bloomington, Indiana: Eurolingua. 2,800 languages listed. **Part I.** Languages of the World in Decreasing Order of the Speaker Numbers with Phyletic and Geographic Identifications. 1986, 136 pp. ISBN 0-931922-24-0. Price (Paper): US $26.00. – **Part II.** Languages of the World Arranged According to Linguistic Phyla with Speaker Number Data. 1987, 112 pp. ISBN 0-931922-27.5. Price (Paper): US $26.00. – **Part III.** Languages of the World Arranged According to Continents and Countries. 1988, 136 pp. ISBN 0-931922-31-3. Price (Paper): US $26.00. – **Part IV.** Alphabetical Index of the Languages of the World. 1988, 142 pp. ISBN 0-931922-32-1. Price (Paper): US $26.00.

Vol. 7 Roman Stopa: **Die Schnalze.** Ihre Natur, Entwicklung und Ursprung. The Clicks. Their Character, Development, and Origin. With an Introduction by Gyula Décsy. Bloomington, Indiana: Eurolingua 1987, 213 pp. ISBN 0-931922-25-9. Price (Paper): US $26.00.

Vol. 8 Gyula Décsy: **A Select Catalog of Language Universals.** Bloomington, Indiana: Eurolingua 1988, 132 pp. ISBN 0-931922-29-1. Price (Paper): US $26.00.

Further volumes in preparation. Orders should be sent to **Eurolingua,** P.O. Box 101, Bloomington, IN 47402-0101, U.S.A. Eurolingua is now the distributor of many fine European publications. Ask for detailed catalog. No charge for shipping on prepaid orders.

In preparation: John Lotz: Das Ungarische Sprachsystem. The Hungarian Language System. Reprint of the 1939 German edition. With an introduction by Thomas A. Sebeok. 304 pp. Eurolingua 1988. ISBN 931922-33-X. *Opus magnum* of the late John Lotz, Professor of Columbia University in the City of New York and of the University of Stockholm/Sweden.

ADVERTISEMENT

ARCADIA BIBLIOGRAPHICA VIRORUM ERUDITORUM

ISSN 0195-7163

Fasciculus 1	*Karl Heinrich Menges Bibliographie.* Compiled by Steven E. Hegaard. 1979. 57 pp. ISBN 3-447-01835/6. US $15.
Fasciculus 2	*Alo Raun Bibliography.* Compiled by Gustav Bayerle. 1980. 29 pp. ISBN 0-931922-02-X. US $15.
Fasciculus 3	*Erich Kunze Bibliographie.* Mit Unterstützung der Friedrich-Ebert-Stiftung. 1980. 33 pp. ISBN 0-931922-07-0. US $15.
Fasciculus 4	*Felix Johannes Oinas Bibliography.* Compiled by R. F. Feldstein. 1981. 51 pp. ISBN 0-931922-03-8. US $15.
Fasciculus 5	*Helmut Hoffman Bibliography.* Compiled by Michael Walter. 1982. 28 pp. ISBN 0-931922-13-5. US $15.
Fasciculus 6	*Fred W. Householder Bibliography.* Compiled by Salman H. Al-Ani. 1984. 37 pp. ISBN 0-931922-16-X. US $15.
Fasciculus 7	*Ferenc Fabricius-Kovács Bibliography.* With support of the Hungarian Cultural Association of Chicago. Compiled by Miklós Kontra. 1984. 41 pp. ISBN 0-931922-18-6. US $15.
Fasciculus 8	*György Lakó Bibliographie.* Zusammengestellt von Edith Vértes. 1985. 37 pp. ISBN 0-931922-21-6. US $15.
Fasciculus 9	*Denis Sinor Bibliography.* Compiled by Ruth I. Meserve. 1986. 65 pp. ISBN 0-931922-12-7. US $15.
Fasciculus 10	*Kenneth Lee Pike Bibliography.* Compiled by Ruth M. Brend. With contributions by A. L. Becker, R. M. Brend, J. Costollo, B. H. Fries, S. Lamb, R. E. Longacre, and A. Makkai. 1987. ISBN 0-931922-28-3. US $16.

Distributed by:

EUROLINGUA
P.O. Box 101
Bloomington, IN 47402-0101
USA